The Journal of
NERVOUS and MENTAL DISEASE

VOL. 128, NO. 2
February, 1959
SERIAL NO. 915

THE RELATION BETWEEN NERVE FIBER SIZE AND SENSORY MODALITY: PHYLOGENETIC IMPLICATIONS OF THE AFFERENT INNERVATION OF CORTEX

GEORGE H. BISHOP, PH.D.[1]

INTRODUCTION

It has been proposed (Dyson, 1958) that in physics, and perhaps in science in general, a new idea can be evolved not oftener than once every thirty years, and that if it doesn't then sound crazy, it is probably worthless. It is just over thirty years since Gasser and Erlanger analyzed the relations of nerve fiber size to fiber properties of conduction rate, electrical threshold etc. Following this, numerous investigators made various attempts to relate fiber size groupings to body functions, particularly to modalities of senses mediated. As in other branches of science, the further testing of these relations, together with industrious pursuit of experiments inspired by the fundamental observations, has preoccupied those interested in the subject. This preoccupation seems to have occluded the possibility of a novel or even a different approach to the significance of fiber size groupings, not to say of fiber size in itself. The proposal developed here practically forced itself on the writer's reluctant attention when he stumbled on the observation that stimulation only of the larger fibers in the optic nerve of the cat caused a response in visual cortex. The smallest fibers innervate the phylogenetically older structure, the tectum mesencephali. That was five years ago; it took three of those years to think of examining the sizes of fibers in the cortex itself. The results have led to a reconsideration of the organization and relations of the sensory systems in general.

We have become accustomed to looking at nervous pathways in cross-sections, involving only two dimensions. The third and essential one for the present discussion is along the neuraxis, and a phylogenetic time dimension might be considered as a fourth. If certain imaginary values appear in its consideration, their employment is not without respectable precedent.

It is proposed here, first, that the fiber size groupings in afferent nerves, which in the somaesthetic system appear to relate to modalities of senses mediated, have also a more general significance in relation to the phylogeny of all the sensory systems projecting to cortex. Second, it is proposed that the higher centers have successively acquired fiber components of successively larger size during their evolutionary development. This can be correlated with the increasingly complex organization and functioning of the cortex. A reexamination is made of the ascending paths in terms of the sizes of the fibers of which they consist, and a reconsideration is proposed of the relation of fiber size primarily to phylogeny rather than to modality.

Much of the information available at present serves only as circumstantial evidence with respect to the above proposals.

[1] Division of Neurology, Laboratory of Neurophysiology and Beaumont-May Institute, Washington University School of Medicine, St. Louis, Missouri. This work was supported in part by a grant from the Supreme Council, 33rd° Scottish Rite, Northern Masonic Jurisdiction U.S.A. through the National Association for Mental Health; and in part by contract 816 (03) between Washington University and the Office of Naval Research.

Its selection, presentation and organization toward an understanding of sensory processes necessarily reflects the writer's bias. Any conclusions here drawn may then be taken as an invitation to the reader to draw more appropriate ones from the data here assembled.

To avoid making repeated qualifications and exceptions to statements about the paths of afferent information, we will employ hereafter the adjective "sensory" in a most general sense. This term and its associated terminology denoting quality, modality, etc. such as pain, touch, temperature, vision, will refer not only to the sensations, but to those *paths* over which impulses transmitted may result in sensation in conscious normal human subjects. Such terminology will be employed even when referring to the lower animals of whose sensations we can only make a sympathetic surmise. Other paths which are afferent but not known to mediate sensations in the human subject will be specified simply as afferent. The latter term usually applies to paths both sensory and non-sensory in their effects.

STATUS OF THE PROBLEM

The perennial problem of the mediation of impulses resulting in painful and other sensations via specific fiber size groups in nerves has recently been illuminated by investigations of the sensory pathways ascending from the spinal cord. This has occurred even before a previous concept—that somaesthetic sensory modalities were to some extent correlated with fiber size groups in peripheral nerves—had gained either general acceptance or conclusive disproof. This disputed concept derives more specifically from the work of Gasser and Erlanger (11) and of Heinbecker, Bishop and O'Leary (14) but is based also on previous work of Goldscheider, von Frey and others of their contemporaries. Early interpretations have been amplified, criticized and occasionally supported by a multitude of more recent workers. Since the present venture proposes to amplify that concept rather than to review the controversial literature, the writer's personal version of the status of this problem up to 1954 will first be stated.

1. Aside from their relations to sensory modality, the different properties of fibers of different sizes serve as useful criteria for identification of paths through the central nervous system. The significant properties may be summarized as follows. The larger the fiber, the faster the conduction of its impulse and the lower its threshold for electrical excitation. In many nerves and central tracts the fibers occur in size groups, such that there are large numbers within certain limits, and fewer between these maxima. Thus when all are stimulated synchronously the action currents after conduction sum to successive maxima, one for each successive size group. The largest fibers of each group have a conduction rate roughly twice that of fibers of the next smaller fiber group, and a threshold of about half. Such groups are often found to have functional connotations in other relations than to sensory modality. The fiber size distribution of a nerve or pathway is referred to as its fiber spectrum,

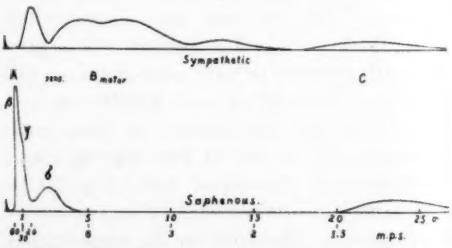


FIG. 1. Diagram of action currents of two types of nerves as recorded after conduction following single shock stimuli. The larger fibers conduct more rapidly and their action currents therefore appear earlier at the recording electrode. The saphenous, a sensory nerve, lacks the B component of myelinated autonomic motor fibers, and the alpha component of muscle nerves. The sympathetic, a mixed nerve, lacks some of the larger sensory fibers of the saphenous. The maxima of the curves for either nerve indicate the presence of concentrations of fibers in the nerve size spectrum of certain size ranges. The largest fibers of the saphenous of the cat are 17 mu in diameter, the smallest (unmyelinated) less than 1 mu.

whether fibers accumulate into specific aggregates or vary more uniformly in size-number relation. In general one may speak of fiber size ranges or regions of the spectrum.

2. In general the larger sensory fibers mediate touch discrimination, pressure, muscle sense, tension etc., and are known from Ranson's work (27) to ascend in the dorsal column of the cord in mammals. After synapse in the gracile and cuneate nuclei the postsynaptic fibers decussate to form the medial lemniscus. Practically all of these (25) pass without further synapse to the ventral nucleus of the thalamus, from which by another synapse the path projects to the somaesthetic area of cortex. The dorsal column fibers constitute the alpha and beta groups of the A fiber afferent range. Most of the alpha fibers terminate in Clark's column from which arises the dorsal spinocerebellar tract (20). The beta group then constitutes or includes the definitive sensory component of the dorsal column. The sensory loss following dorsal column damage serves as evidence for the above correlation between fiber size and sensation.

3. The small myelinated gamma and delta fibers mediate touch also, plus pain and temperature, in an order not readily determinable, since the functional ranges overlap with respect to fiber size and threshold. The exact locus of their cord synapses is not accurately known, but appears to include the substantia gelatinosa. Postsynaptic fibers cross in the ventral white commissure to ascend in the anterolateral column of the cord. Certain of these fibers make up the primate spinothalamic tract, some fibers of which are known to terminate in the ventral nucleus of the thalamus, entering it with the medial lemniscus bundle.

4. The C or unmyelinated sensory fibers mediate a second type of pain, delayed or burning in character as contrasted with early, *i.e.* short-latency pricking pain. The existence of a second pain has been reported frequently, and about as frequently denied. Delayed pain is most clearly demonstrated (by its advocates) in the absence of pricking pain, whether this absence is accomplished by disease or by experimental block of all but unmyelinated fibers. It has been proposed that its efficacy or intensity is masked or reduced by the presence of pricking pain, or by the functional presence of the delta fiber component, even when this is not stimulated. The C fibers take the course of the delta fibers at least to the substantia gelationosa of the spinal cord, and their secondary path accompanies the delta in the spinothalamic tract, though whether the path continues as C fibers, and whether it terminates in the ventral thalamic nucleus or has further synapses along its course has not been investigated until recently. From the latency of the delayed pain it has been inferred that the conduction rate probably indicates a pathway consisting of very small fibers, and one which probably contains several synapses.

Items 2, 3 and 4, comprise a statement of certain current points of view, that the writer now belives give a limited and in some respects erroneous picture of the significance of fiber size in relation to sensation mediated. The following account of recent work may justify a revision of the correlation of fiber size with sensation.

RECENT INVESTIGATIONS

Since a major aspect of any discussion of bodily sensation will involve the problem of pain, a preliminary remark is in order as to how the term is used. Hall (13) has recently proposed the dichotomy of pain and suffering; suffering can be experienced without pain, and pain may not involve suffering, unless severe and presented under conditions arousing anxiety or other emotional disturbance. The potentially limited painful sensory experience below the level of suffering is that which will be referred to here as painful *sensation*. Though not always taken into account, the same disturbing factors so obviously conditioning the estimation of the

quality of painful sensations also affect, if to a lesser degree, the reporting of any sensation. It would be fortunate indeed if the physiologist, attempting to analyze the afferent paths of the nervous system, could disregard subjective reactions; but in fact clinical and experimental findings in the conscious human subject are invaluable in defining even the courses of the anatomical paths by which sensations are mediated. Such findings may indeed be more significant to the problem of the *sensory* pathways than the observation of lower animal behavior, which is only secondarily applicable, by way of human inference of what the animal may experience.

Sensory vs. afferent paths

This situation emphasizes the fact that the general problem is the analysis of *afferent* paths, not only of their sensory components, however valuable the information derived from the activity of the latter may be as a cue to the functioning of the former. Each afferent path and in fact each fiber-size component of that path sends its output to many levels of the nervous system. One often speaks of collaterals of a path, as if the lower levels of the nervous system were innervated by collateral branches of the fibers which primarily terminated at higher levels. Rather, even when a given fiber may innervate two levels, the lower level is the more primitive terminus and a branch extending to a higher level should be the collateral. In any event there are clear cases in which different fibers, though of common size and from common origins at the spinal level, innervate one *or* another level. For instance large fibers of the dorsal column terminating in medulla and brainstem are distinct from those synapsing in the cuneate nucleus and postsynaptically activating the thalamus. The same applies to the lateral column afferents, a majority of which terminate below the thalamus as separate fibers from those of the spinothalamic tract. From the phylogenetic point of view in any case, an account which starts from a consideration of the sensory paths to consciousness must shift emphasis to a tracing of successive additions to the paths of lower forms, additions which have occurred during the evolution of higher forms. It will be argued in what follows that the additions have been of two types; first, additions of fibers similar to those already present, but bypassing the primitive levels of termination to reach the newly developed higher centers, and second, invasion of these higher levels by successively larger fibers as these centers become more highly differentiated in function. The advantage in approaching this question through the analysis of *sensory* paths is that the thalamus and cortex in which they terminate have a known phylogenetic history in the vertebrates, and some of their sensory inputs are specific and accessible to study.

The problem of identification of pain paths

Two comments are apropos regarding the identification of the sensations, and particularly of pain, as the result of activity in specific groups of fibers. First, not all afferent fibers ascending in either dorsal or lateral columns of the cord reach the thalamus. Petren (26) reported in 1910 that after degeneration in man following cord section about 75 per cent of the presumed spinothalamic bundle at the cord level terminated in the reticular formation of medulla and brainstem. A similar termination in the reticular formation of many fibers ascending from the cord has been emphasized by Herrick (15) both in mammals and in the lower vertebrates. At least four attempts to demonstrate a spinothalamic tract in the cat by Marchi technique have failed to reveal fibers passing without synapse to the thalamus. This and other considerations have led various authors to infer that the pain pathway ascends through the reticular substance by a series of synaptic links. No pain fol-

lows stimulation of the somaesthetic area of cortex nor is pain usually interfered with after the removal of that area. These circumstances, together with the phenomenon of the thalamic syndrome, have raised doubts that even in the primates the pain path enters the cortical projection area via the ventral nucleus. Instead, the inference has been made that it may terminate or project primarily in other regions of the thalamus itself.

Second, patients frequently report painful sensations that cannot be referred to the presumed appropriate fiber categories. It is also true that any strong experimental stimulation to the body periphery can induce pain. These facts have led to the inference that the result of the excessive stimulation of *any sensory fibers* is painful. Such inferences have not been adequately verified. While it must be conceded that pain to the subject is what he says hurts, in the writer's experience such reports are often made by patients in states of anxiety etc. due to apprehension of associated sequelae of injury or disease, or to fear of the experimental procedure. One of the most common instances involves the interpretation of dysaesthesia or tingle as painful. When tingle is arbitrarily induced, as it can be by repetitive stimulation of only the largest fibers of lowest threshold in a peripheral nerve, the patient recognizes the sensation as that causing his distress, but under these conditions no longer calls it painful. In any case it appears necessary to differentiate, in the subject's experience of sensation, between a relatively objective report of the quality of sensation as such and the more emotional interpretation of what the sensation implies. Hall's (1957) and other proposed dichotomies of painful sensation vs. suffering, reaction, affect, etc. are pertinent here. But even when confusion exists in the subject's reports of his casual experiences, it is still possible to arrive at a clear differentiation between the sequelae of *experimentally* stimulating different groups of fibers.

RECENT INFORMATION ON THE COURSES OF THE ASCENDING SPINAL TRACTS

In what immediately follows, certain pertinent results of anatomical investigation and of physiological experiment will be presented insofar as physiological recording has given evidence as to the anatomical course of sensory pathways. The emphasis will be less on modality of sensation mediated, than on the paths by which sensory impulses reach higher centers.

From the point of view of sensory pathway differentiation, a premonitory finding was reported in 1952 by Magoun and his coworkers (22), that following stimulation of a locus in the upper brainstem of cats, a change occurred in the record of spontaneous activity of the cortex. This change is similar to that observed in the electroencephalogram when a subject receives any sensory stimulus, or even engages in mental activity. The alpha rhythm is reduced or abolished, and replaced by a low voltage discharge which often approaches a flat baseline. The result is referred to as the "activation pattern." To produce such a record following brainstem stimulation Magoun found it necessary to stimulate electrically at a frequency of 30/sec. or more. Below this stimulus frequency the cortex shows repetitive responses, if any occur, at the frequency of the stimulus, superposed on a more or less altered spontaneous pattern. With stimulation of a peripheral nerve and central recording, a response was detected in the mesial reticular substance of the rostral brainstem, caudal and ventral to the centrum medianum nucleus (21). This response and that recorded from the ventral nucleus of the thalamus were compared for the same peripheral stimulus. The midbrain response developed later, but was thought to have approximately the same threshold in terms of the peripheral nerve stimulus. For this and other reasons it was inferred that collaterals of the classically known direct sensory paths served a function which is to activate the cortex diffusely.

Later, Collins and O'Leary (8) re-examined the mesial brainstem and ventral nucleus responses to peripheral sensory nerve stimulation in the cat. They decided, first, that the response from the ventral tegmentum required a stronger peripheral nerve stimulus than did that from the ventral nucleus of the thalamus, and second, that a mesial thalamic response followed the medial brainstem response presumably after synaptic passage. Careful monitoring of the peripheral nerve response while recording the central responses to the same stimulus indicated that the large low threshold dorsal column fibers activated chiefly the ventral nucleus response, while smaller higher threshold lateral column fibers activated the mesial brainstem response. In fact the former corresponded to stimulation of the beta nerve axon group, the latter to gamma plus delta. On transecting the lateral column of the cord the medial responses dropped out; on transecting the dorsal column the ventral nucleus response was reduced to a trace. From these findings it could be concluded that the medial brainstem responses followed the activation, not of a path collateral to the main sensory paths, but that of the lateral column sensory tract itself. In the cat this presumably includes the gamma and delta path for touch, pain and temperature. The possible contribution of C fiber activity to these responses was not definitely established.

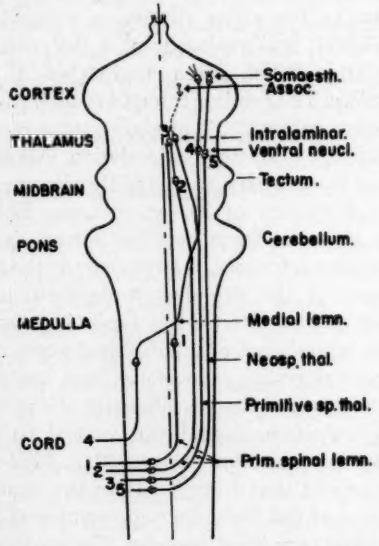


FIG. 2. Schematic diagram of the central nervous system including five sensory paths ascending from the spinal cord. Only two (4 and 5) project to cortex through the ventral nucleus of the thalamus. These are present only in mammals. The other three constitute the spinal lemniscus system, present in both premammals and mammals, with terminations at the medulla, brainstem and intralaminar nuclei of the thalamus respectively.

The central course of the "spinothalamic" fibers has recently been reinvestigated by Mehler (23) employing the Nauta technique for revealing axon degeneration. All previous investigations had been carried out by the Marchi method which stains degenerated small fibers poorly, and unmyelinated fibers not at all. This limitation may account for previous failures to find by anatomical techniques the medial path in mammals identified physiologically by Collins and O'Leary. Mehler investigated a series of animals ranging from the opossum to the chimpanzee. He found in all a medially coursing path of small fibers, some of which pass without synapse from spinal cord to intralaminar thalamic nuclei. No degeneration was observed in the centrum medianum, although this medial tract passes near the centrum medianum, which may account for Magoun's and Collins and O'Leary's use of this nucleus as a landmark. This medial path from the lateral column of the cord takes approximately the course of the paleospinothalamic tract of the premammalian vertebrates discussed in detail by Herrick. Mehler found that with ascent of the phylogenetic scale this medial tract became relatively less prominent, while an increasing number of fibers passed to the ventral nucleus, forming the classical or neospinothalamic tract so well known in the primates. It seems probable, although it has not yet been demonstrated, that the pain

and temperature paths, in primates as in lower forms take a medial course through the higher brainstem, and include both the through tract traced by Mehler and the component traced by Collins and O'Leary, which is relayed in the medial brainstem to the intralaminar thalamic region. This region of the thalamus of mammals has not been recognized previously as a sensory receptive area. When this thalamic region is stimulated electrically, the cortical recruiting responses first described in 1942 by Morison and Dempsey (24) are recorded.

A difficulty is involved in this interpretation; the intralaminar nuclei do not project immediately to any area of neocortex as is indicated by the absence of retrograde degeneration after cortical removal. The recruiting responses following stimulation here and recorded over large areas of cortex have been accounted for by mediation via relays to other thalamic nuclei which do project diffusely to neocortex. The same path could be invoked for diffuse mediation of pain from the intralaminar region. An alternate or perhaps parallel possibility for the sensory pain path termination in the thalamus is provided by the work of Woolsey and Rose (in press) on thalamocortical relations in the monkey. They find, as do Nauta and Kuyper (25) that not all the fibers of the classical spinothalamic tract terminate within the ventral nucleus itself. A majority of them terminate just ventral and posterior to this nucleus. Retrograde degeneration is found in this region of thalamus following removal of the secondary somaesthetic area of cortex. This area in the human, angular and supramarginal gyri, corresponds to part of the area in cat and monkey occupied by the secondary somaesthetic area. Interference with or loss of pain sensation has been reported (28) in several human cases after injury or surgical interference in this region.

It may be of interest that Morison and Dempsey, in the one instance in which they reported an "augmenting" response from stimulation in the region of the ventral nucleus, separate from involvement and confusion with the specific sensory response of the somaesthetic area, stated that their stimulus was applied in the margin of the ventral nucleus. The response lacked the characteristics of the definitive pattern of response to sensory nerve stimulation, and otherwise the augmenting responses have almost identical form and other characteristics of the recruiting responses elicited by stimulation of the intralaminar and other regions of the thalamus.

Collins and Randt (1958) undertook to follow the path postsynaptic to the sensory C fibers upward through the nervous system, by means of electrical recording of action currents. Using strong alternating currents they blocked all but C fibers in the nerve (since the latter are the most resistant to this procedure) and found that repetitive stimulation of the C fibers caused a strong repetitive response in the medial reticular region of the medulla. The record appears to be that of a postsynaptic nuclear activity, persisting for some time after the stimulus is discontinued. Nauta and Kuypers (1958) have recently reported on earlier and more recent evidence of paths from medulla and brainstem to certain intralaminar thalamic nuclei and to the subthalamus, judged to constitute relays from afferent fibers possibly mediating pain. Bowsher (5) in a recent review of the reticular connections also describes reticulothalamic paths that might relay this activity upward. Bowsher in fact accounts for the long latency of responses to pain as compared to those for touch as being due to the synaptic times for successive relays upward through a sequence of short segments of path; an idea which was early suggested by observations that many fibers from the lateral column terminated in medulla and brainstem. Mehler's findings may make this assumption unnecessary, but not necessarily in error.

Still more recently Collins and Randt (personal communication) have traced the activity due to C fiber stimulation at least

as far as the midbrain level previously investigated by Collins and O'Leary. Since pain results from stimulation of C fibers, this again points to a medial destination for at least one component of the pain pathway, and there is as yet no indication of C fiber activity in the region of the ventral nucleus. In the present, if not final, analysis of afferent paths to thalamus, Nauta and Kuypers in effect describe a fan of fibers radiating from the ascending lateral cord tracts to a thalamic region extending from the intralaminar nuclei to the ventral nucleus, though concentrating their terminals in the medial and lateral positions of this basal region of thalamus. One might infer that touch and proprioceptive paths are distributed toward the lateral aspect of this region, centering in the ventral nucleus, and pain and temperature toward the medial region. A final designation of certain parts of this distribution, direct or relayed, as specifically sensory is still precarious except for the ventral nucleus projection system. Also the identification within this system of the exact homologue of the premammalian (paleospinothalamic) tract (see Figure 3) must wait on further investigaton of functional as well as anatomical distribution in the afferent somaesthetic system. As to the afferent paths, more inclusive than the sensory, the general distribution seems to be that the larger fibers terminate more laterally in the thalamus, the smaller fibers more medially.

PHYSIOLOGICAL ANALYSIS OF SENSORY PATHS

The gamma-delta range of small-myelinated lateral column afferents. That the delta group of small myelinated fibers mediates pricking pain has been reported by many authors with some doubts expressed that it is strictly confined to this range. That the pain path ascends the lateral column of the cord is a neurological finding upon which is based the often completely successful operation of lateral cordotomy. Mediation of temperature sensation has been assigned to the lateral column path by many authors going back to von Frey and Goldscheider. These sensations have been related to the small myelinated fiber path by differential anaesthesia and other blocking procedures (14). Temperature lies just below pain within the delta range. Evidence that a component of itch is representated in the delta range has been presented by Bishop (2). Thus in this group of fibers—the small-myelinated—touch, itch, pain and temperature are represented. The postsynaptic paths from these gamma and delta fibers pass up the lateral column (Collins and O'Leary) as do the unmyelinated axons, and many take the course of the paleospinothalamic path to brainstem relay and to thalamus direct (Figure 2). Some of these fibers deviate from this path to constitute the more laterally coursing neospinothalamic tract, the classical spinothalamic of man, terminating in or near the ventral nucleus.

The C component of unmelinated fibers. In the human subject after block of arm or leg by cuff pressure of 250 mm. Hg or more for 25 to 40 minutes, all touch and pricking pain is lost (19). Stimulation of the skin, or more potently, stimulation of a nerve below the block by electric shocks then causes excruciating burning pain. In animals such pressure has been shown to block delta fibers leaving the C fibers to conduct undisturbed. Shocks to the nerve summate strongly at 3 to 5 per second, in contrast to normal pricking pain which requires a higher frequency for smooth summation. Such pressure block of pricking pain leaves unchanged the burning pain induced in severe tabetics by strong skin stimulation. On the other hand dilute novocaine blocks C fibers in excised nerves before delta fibers. Novocaine injected around a nerve blocks burning pain in a concentration below that required for block of pricking pain. By this drug, the pain of inflamation of deep tissues, and the "protopathic" pain of Head and Rivers detectable at the margin of a de-

nervated area can be blocked, leaving pricking pain intact. Thus the pain that can be blocked is similar to the delayed pain reported to be induced by various procedures. A number of other sensations can be correlated with C fiber activity.

Graham, Goodell, and Wolff (1957 and references cited there; Bickford 1938) reported itch assignable to C fiber stimulation after Bishop (2) had concluded that itch was induced by weak stimulation of delta fibers. Recently Shelley and Arthur (29) also report itch assignable to C fiber activity, stimulated by techniques differential for its production. This has been confirmed (Landau and Bishop, 1958) after pressure block of delta fiber pain, when using Shelley and Arthur's technique of intradermal stimulation and also by use of cowhage (itch powder). The threshold for itch rises after delta fiber block indicating that delta fibers also contribute to the effect, as was inferred by Bishop and by Graham *et al.* Warmth and cold can also be perceived after delta block (19).

Finally by the use of an ingeniously devised experiment, Douglas and Ritchie (10) have concluded that stroking the cat's fur with a roll of gauze, which cannot be conceived to arouse pain, sets up a very pronounced activity in the saphenous nerve. The frequency of response and rate of conduction of a major component of this activity indicate that C fibers respond to tactile stimuli. Landau and Bishop have not been able to detect a tactile *sensation* in human subjects from skin stimulation after myelinated fiber block. Either the human differs from the cat in this respect, or the C fibers activated by tactile stimuli are not represented in consciousness. Douglas and Ritchie, using the cat, however, have demonstrated tactile afferents in the C group, whatever sensory connotation this may prove to have. Thus within the C range are represented a full complement of somaesthetic modalities; tactile, itch, pain and temperature, duplicating the modalities mediated in the gamma-delta range. Correlated with the results are those of Collins and Randt tracing C fiber activity to a synaptic relay in the medulla, and their more recent recording (unpublished results) of responses to C fiber stimulation in the higher brainstem. The postsynaptic path for mediation of these C modalities can be inferred to take a medial course above the spinal level through such a reticular path as that represented by Nauta and Kuyper (1958) as a medial reticulothalamic tract (Figure 2). Persons congenitally lacking all pain appreciation still discriminate sharp from dull in skin contact. Mehler and Nauta (personal communciation) suggest that a neo-spinothalamic component may contribute a quality of sharpness or acuteness, an "epicritic" quality, to pain sensation which is summed centrally with a "protopathic" component mediated over the paleo-spinothalamic path to the intralaminar region of the thalamus. The small-myelinated fiber path is at least distinct from the C fiber path above the spinal level.

Large fiber dorsal column sensory afferents. The dorsal column large fiber component is known to mediate touch, tension, muscle sense, etc., on the basis of neurological findings after damage to the dorsal cord. Pain and temperature are not lost after such damage, and no evidence that they are mediated by dorsal column fibers has been reported. Lloyd and McIntyre (20) have demonstrated that in the cat, group I fibers of deep nerves (A, alpha) pass to Clark's column for relay to the neo- or dorsal spinocerebellar tract; and these do not convey activity to the conscious level. Correspondingly, no alpha wave appears in the saphenous and other cutaneous nerves. This leaves the beta size range alone to relay through the medial lemniscus. The senses mediated appear to be those for fine discrimination and proprioception; two point touch discrimination, vibratory sense, recognition of figure writing on the skin, tension and joint sense, and position in

space. Thus not all the modalities represented in the lower size ranges are present; but tactile functions are represented profusely in this third system of fibers. Fibers mediating the sense of pain are conspicuously absent here as contrasted to the smaller fiber size ranges. A finely discriminative pain function appropriate to the dorsal column system is perhaps yet to be evolved; it may be anticipated indeed by the painful experience ("suffering") often associated with non-painful sensations.

PHYLOGENETIC RELATIONS OF THE AFFERENT PATHS

The lemniscus systems. The foregoing account comprises a physiological and histological analysis of the main somaesthetic sensory paths in mammals. It indicates that the different fiber size groups of peripheral nerves segregate into separate paths at higher levels, and project to different regions of higher centers. The rationale of this arrangement becomes evident in the history of their evolutionary development. The origin of these paths in lower forms has been studied extensively by Herrick (1948 and references there noted) and recently summarized with particular reference to the comparison between the premammalian vertebrates and the mammals (18). The following conclusions from such studies are pertinent to further analysis of the sensory systems.

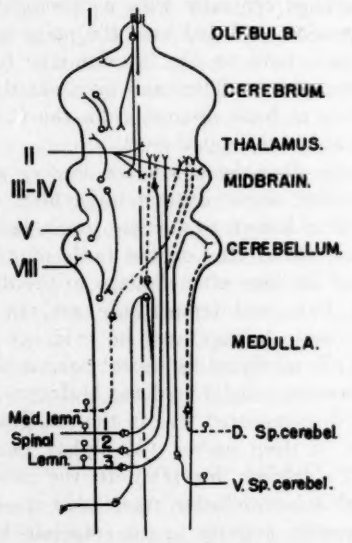


FIG. 3. Scheme presenting the courses of the spinal lemniscus components 1, 2, 3, and their primary projection levels at medulla, midbrain and thalamus respectively. The paths in full lines are present below the mammals. Those in dash lines are added as the higher centers of mammals develop. The primitive olfactory and visual afferent systems are indicated by fine full lines on the left half of the diagram. A few of the inter-level connections are also indicated by curved lines at left.

1. In lower vertebrates many spinal afferents synapse homolaterally, and postsynaptic axons cross the cord to ascend in the spinal lemniscus (Figure 3). Some of these fibers terminate in the medulla to constitute the afferent limb of a spinal total-pattern activity. However, in Amblystoma, for instance, some such ascending axons of the secondary path continue to brainstem, tectum and thalamus. The thalamic component of this ascending bundle comprises the *paleospinothalamic tract* which may be the precursor of the through tract demonstrated by Mehler, projecting in mammals to the medial thalamic region.

2. Fibers which project through this primitive spinal lemniscus to the upper brainstem should represent the precursor of the path recorded there by Collins and O'Leary. Since stimulation in this brainstem region produces responses at the cortical level, this might be accomplished by relay through the region of thalamus in which Mehler's paleospinothalamic tract in mammals terminates. Accompanying these extensions of the local spinal lemniscus system is the primitive homolateral spinocerebellar system, the precursor of the mammalian ventral spinocerebellar tract.

3. No true medial lemniscus is present below the mammals, although in the spinal cord there is a dorsal column through which

fibers pass upward to terminate in the reticular formation. In mammals a new system of fibers parallel to or differentiated out of these arises in the cuneate and gracile nuclei. This bundle crosses the midline in the medulla. All the fibers from these nuclei pass to the ventral nucleus of the thalamus, and relay to the somaesthetic projection area of cortex. This sequence of new structures in the mammal, constituting a direct path to neocortex, bypasses in a sense the paleospinothalamic tract, terminating more medially in the thalamus. Many of the largest fibers of the dorsal column (alpha fibers of peripheral nerve) relay in Clark's column to form the neocerebellar tract; beta fibers chiefly supply the medial lemniscus.

4. Again only in mammals, some fibers from the lateral column digress from the medial paleospinothalamic path at the level of the inferior colliculus, and join or pass parallel to the medial lemniscus bundle. Some relay in the ventral nucleus to project certain of the lateral column senses to somaesthetic cortex. Since pain seems not to be projected to this area, it may be inferred that touch fibers are chiefly those diverted from the older system.

5. Still another sensory route is that of Collins and Randt consisting of unmyelinated fibers which relay in the cord and ascend in the lateral column to terminate in the medulla. From here it is presumably relayed upward, through one or more synaptic levels, as a medial reticulothalamic tract.

These five afferent paths can all be activated by stimuli applied to peripheral nerves, where thresholds and conduction rates characterizing their fiber size ranges serve to identify their origins. In conscious human subjects and in correlated animal experiments the modalities of sensation mediated over the fibers of different sizes can then be related to their courses through the central nervous system.

Functional connotations of the phylogenetic sequence. Of the five paths outlined (Figure 2) only two, those developed exclusively in mammals, project to the somaesthetic area of neocortex via the ventral nucleus relay. They mediate only tactile and proprioceptive modalities, lacking pain and temperature. This means that the somaesthetic projection area of cortex is chiefly devoted to the analysis of only part of the sensory periphery, although it must be considered probable that other sensory modalities are somehow represented at the cortical level. Deferring this consideration for the present two questions are worth raising. First, of those sensations mediated by more than one path, *e.g.* touch in three, pain and temperature in two, what is the significance of the duplication? Second, since all the main sensory systems, somaesthetic, auditory and visual, have many features in common, do the direct cortical projection paths of each involve only the larger fiber components from the periphery, mediating only certain types of sensory information?

The qualities of sensation within one modality

Concerning the first question, the qualitative characteristics of the two varieties of pain and of the types of tactile sensations experienced permit at least one generalization; that the sensory information mediated by the larger fibers, over central paths more recently acquired and more directly projected to higher centers, is more precise and specific than is that mediated by the more primitive systems. For instance, C fiber pain is diffuse and poorly localized, as contrasted with delta fiber pricking pain, which exhibits more accurate localization and fairly good two-point discrimination. Nothing is known concerning possible C fiber tactile sensation, but fine two-point tactile discrimination, perception of figures drawn on the palm, accurate localization of the limbs in space and vibratory sense are lost

or impaired after dorsal column damage. This difference in the refinement of interpretation of the periphery is of course not assignable alone to peripheral mechanisms; the central connections after all determine the final result. But at least the differences found are consistent with the differences in higher pathway contacts with medulla, thalamus and cortex, and add evidence that the more recent duplication of primitive functions within a given sensory modality accomplishes an advance in the functional competence of the whole sensory apparatus.

This thesis implies that the smaller fiber systems not only make connections to their respective higher centers earlier in phylogeny than do the larger fiber systsms, but that smaller fibers are more primitive in general than larger fibers. Some evidence can be adduced to this effect. A phylogenetic differentiation between unmyelinated and myelinated fibers as such is lost in the mists of even greater antiquity than is the evidence for successively developed pathways. Large fibers are present in the spinal systems of lower forms, though the higher and more recently developed centers of these forms lack them as if they had not yet extended the large fiber paths to higher centers in the present stages of their evolution. In the turtle cortex unmyelinated fibers predominate although relatively small myelinated fibers are present. In some lower mammals as in the premammals many unmyelinated fibers are present in the optic nerve, which are replaced by small myelinated fibers in the cat, human and a number of other forms (6). As has been noted above, the unmyelinated path in the somaesthetic system of the cat terminates at a lower level of the nervous system than do the myelinated paths. The idea thus seems tenable that the unmyelinated fiber represents a more primitive condition in general; that successively larger fibers have been acquired later even in the older structures of the nervous system; and that fibers of increasing size have successively invaded the higher levels as these became developed.

Comparison of the main sensory systems: modality vs. phylogeny as criteria

As to the second question, the applicability of these considerations to other sensory systems, little is known concerning the auditory apparatus in terms of fiber sizes; certain information is available however concerning the visual system. As a diencephalic structure the eye must have first made its connections with the thalamus as a sensory integrating center, a condition of which the persistent anterior accessory tract may be the evidence. It early sent fibers of the posterior accessory tract to the head of a motor system in the ventral region of the brainstem (16). An obvious extension or addition to this as the tectum became important as a motor control center or integrator consists of the brachium of the superior colliculus. A geniculate nucleus is present below the mammals but without a specific relay to cortex. In the higher mammals the dorsal nucleus of the lateral geniculate body becomes semidetached from the thalamic mass and has long been recognized as the seat of a specific relay to visual projection cortex. These phylogenetic stages are readily recognizable in the cat (3) and here the fiber size relations in the optic tract distribution become significant for comparison with the homologous somaesthetic sensory system arising from cord and medulla.

The fiber size distribution in the optic nerve of the cat has been analyzed both histologically and physiologically. The fibers are all myelinated (6) although in certain lower mammals many unmyelinated fibers are present. The largest fibers are 12 mu in diameter, the smallest not over 0.3 mu, as measured on electron microscope photographs (Bishop, current work). The larger fibers innervate the more recently developed structures. Only the largest 10 per cent or less of a size down to 5 or 6 mu

relay through geniculate to cortex. The next lower sizes also relay in the geniculate, their stimulation resulting in activity not in cortex but in the lateral nucleus of the thalamus; a thalamo-thalamic distribution. A still smaller-sized group innervates the pretectal area for pupillary response to light. The smallest fibers terminate in the colliculus. The presumably still more primitive paths, anterior and posterior accessory tracts, have not been studied in this connection.

These fiber size relations can be checked by electrophysiological recording, employing the established relations between fiber sizes, electrical thresholds and rates of conduction. The optic nerve action current spectrum following stimulation in the eye socket shows two prominent waves recorded at the geniculate level, corresponding to two larger groups of low threshold and rapid conduction, comparable to the beta and gamma fibers of spinal nerves. A remainder of low potential has a higher threshold. This latter is difficult to deal with directly but is readily correlated with large postsynaptic responses in pretectal area and colliculus.

On the basis of this information, certain correspondences between visual and somaesthetic systems are demonstrable. The cortical projection component of the visual system corresponds to the dorsal column-medial lemniscus component of the somaesthetic system, and consists of fibers with properties close to those of the beta group of peripheral nerves. The next lower group should be homologous with the gamma-delta range of peripheral nerves, and also projects in the visual system to thalamus rather than to cortex. The smallest fibers, all myelinated in the cat optic tract but grouped in size about a one mu peak, should correspond to the unmyelinated fibers of peripheral nerves. Their function also corresponds in the sense that they have as their terminus the tectum which became prominent at an early stage of phylogeny as a motor control center, coordinating motor responses to visual stimuli as the medulla did for the spinal system. A few larger optic tract fibers of approximately gamma size range, innervating the tectum are presumably a secondary overlay of the much more numerous small fibers.

The inference drawn from this analysis is that although fiber groups corresponding to and presumably analogous to those of the somaesthetic system all serve one function, visual, the relation is rather obviously to the phylogeny of the visual system rather than to its functional differentiation into modalities. Even if such visual functions as perception of form, color and frequency, analogues of space, tension, vibration, etc. in the dorsal column somaesthetic projective system, are thought of as separate modalities, these visual functions, like the somaesthetic dorsal column senses, do not differ so far as is known with respect to sizes of fibers serving them. No *sensory* response seems to be induced via the tectum in the cat, although in anurans and in birds the thalamus and possibly the tectum may be involved in whatever sensory response these animals may have, since the cortex is poorly developed. In these lower forms the maximal fiber size in the optic nerves is smaller (4–6 mu in the bird) thus lacking the fiber component which in the cat projects to cortex as the principal, and perhaps the only path to visual perception. These smaller fibers in the cat may have lost their sensory function as the higher animals transferred their emphasis for vision to cerebral cortex. In the somaesthetic system by comparison all the fiber size components have retained sensory correlates, with the possible exception of tactile perception in the C fiber system of man. With respect to other than sensory correlates, in visual and somaesthetic systems, fiber size range and distribution to higher levels correspond. From this point of view a functional cor-

relation of fiber size with sensory modality may be little more than coincidental.

PRIMARY PROJECTION LEVELS OF THE SENSORY PATHS.

A third set of relations further illustrates that the various sensory systems show features in common, and similar patterns of organization suggesting parallel courses of phylogenetic development. This appears in terms not only of anatomical paths, but of types of response that occur along these paths. In both optic and somaesthetic systems a volley impulse from an electrical stimulus in the periphery causes responses that can be observed by means of oscillographic recording. These responses are of two general types. The lower synapses in these paths transmit relatively simple volleys; that is, although the postsynaptic elements may each respond more than once to a single stimulus, the relay seems to be chiefly from one group of axons to one group of cells at that locus. A synaptic region is finally encountered where one group of cells after another responds, prolonging the discharge recordable from a given locus. Such a region can be recognized for each of the paths ascending from the different fiber groups of peripheral nerves. By analogy with the concept of cortical projection, we may refer to these synaptic regions as the primary projection levels of these paths, as levels at which not a simple relay but a more complex function is performed.

Relay vs. terminal synaptic regions; target nuclei or projections of phylogenetically successive afferent paths

Such a difference of pattern appears between the relatively simple ventral nucleus response and the more complicated sequence of activity of the somaesthetic cortex, denoting the cortical projection area as the target of the large-fiber paths through the ventral nucleus relay. A similar change of response pattern occurs along the paleospinothalamic path at a lower level, in the intralaminar region of the thalamus. In Collins and O'Leary's records of brainstem and medial thalamic responses to nerve shocks the brainstem response is relatively simple and brief, while postsynaptically a

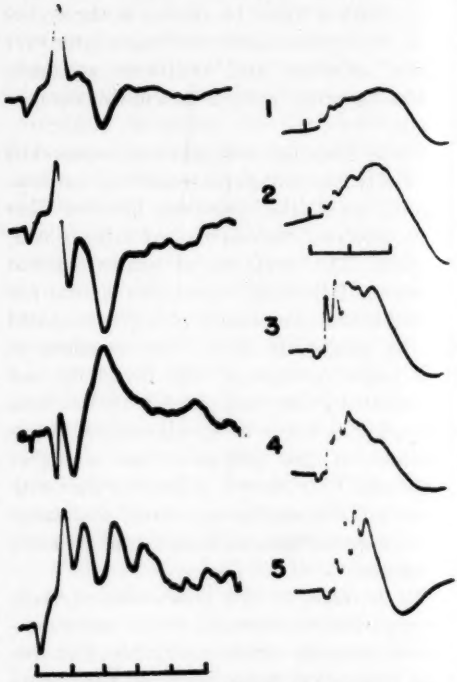


FIG. 4. Responses (left column) at the geniculate and (right) cortical levels following a single stimulus to the optic nerve of the cat. In the upper geniculate record the first spike upward is the presynaptic response, the following spike downward is that of the postsynaptic elements. In the second record the stimulus is increased, adding a few higher threshold, slower conducting fibers of a second group. The third record represents the same events but from recording electrodes in a different position. The fourth record shows the results of a still stronger stimulus, inducing two presynaptic spikes followed by two postsynaptic responses. The cortex response (right) is maximal when only the first spike and its postsynaptic response are activated at the geniculate level. Even a weak stimulus (top) is followed by a complex record, each minor spike representing the discharge of a different group of elements. With increased strength of stimulus the spikes become more prominent but not more numerous, and a negative slow wave follows them.

longer-lasting and complex sequence of spikes signalizes involvement of a complex group of neurones as a target in thalamic and subthalamic regions. In Collins and Randt's experiments following the stimulation of C fibers only, after block of all myelinated components, this appearance of a postsynaptic complex discharge occurs at a still lower medullary level.

From these circumstances the inference may be drawn that the more primitive the sensory path the older and more distal is its primary projection and the less immediate are its relations to cortex. These relations obtain also in the visual system where a geniculate relay response consists of a single spike, while the sequelae in cortex and lateral nucleus of the thalamus are complex and longer lasting. This does not deny that at any synaptic level connections may be made for reflex and other functions. It rather connotes that a heirarchy of primary target organs occurs in those components of paths which lead upwards to consciousness of sensation. The fact that in these ascending bundles the numbers of fibers reaching the thalamus is far smaller than the number leaving the cord indicates that local level connections made by fibers terminating in lower structures along the way are in part at least different fibers and not collaterals of those in the main path to sensation. Apparently in the process of cephalization the extension of each spinal path to higher centers has occurred not only by addition of synaptic links above its primitive terminus, but also by adding more fibers in parallel with the fibers already supplying lower level functions.

Whatever the details of this process may have been, the unmyelinated path has so far as known at present only one relay station, that in the spinal cord (Figure 1) to reach its target in the medulla. The old spinothalamic path in the mammal has some fibers that again have only one relay, in the cord, to thalamus, and others that relay through a second synapse in the medial brainstem to reach the same region of thalamus. The neospinothalamic path and the medial lemniscus have two relays, one in the cord and one in the ventral nucleus, to project to cortex. Insofar as consciousness in higher animals involves cortex, one or more synaptic passages above the primary projection level must be required in the small-fiber paths, as will be considered later in more detail; but conscious recognition of differences in sensation mediated via different paths gives evidence that these paths must converge to a common set of neurones, and the paths appear not to join below this. It would further appear that in some remote time the target organ of each of these more primitive paths constituted the central terminus of its activity, and that extensions of path beyond this initial terminus were added as higher centers became dominant in the course of cephalization. Thus we may think of the unmyelinated path's terminus in the medulla as a most primitive center for organizing spinal activity, only secondarily extended upwards as a sensory path when the higher sensory levels were organized to receive it. This lowest level is bypassed by the small fiber paleospinothalamic path as a more specific and direct route to the thalamus, when the latter became the dominant sensory coordination apparatus. The most recently added paths, medial lemniscus and neospinothalamic tract, constitute similar bypasses to the older thalamic function still served by the paleospinothalamic projection. But this older projection must itself be extended in a sense by another link to reach cortex, to the degree that cortex is involved in the senses mediated.

THE SIZES OF NERVE FIBERS INNERVATING NEOCORTEX

If there has been in fact a successive invasion of higher centers by successively larger fiber groups from spinal levels, a corresponding contribution to higher level functioning should have been evolved on

the basis of more precise or discriminatory information about the periphery. The fiber size distribution in cortex itself might be expected to show corresponding developments to take advantage of this information. Preliminary study of fiber size distribution in the white matter below various regions of cortex indicates that the aquisition of larger fiber groups is one of the characteristics of mammalian *neo*cortex. In the turtle cortex (Orrego, current work) a majority of the fibers are unmyelinated; the remainder are small-myelinated. In the cat and man, fibers up to 13 or 14 mu in diameter are present. In electron microscope photographs of cat and human white matter (Bishop, work in progress) the smallest fibers are myelinated (with exception of so few unmyelinated fibers as to raise suspicions that these could be glial in origin). Many fibers are less than 1 mu in diameter, with a peak in numbers at or below this figure. These small fibers become unmyelinated as they approach the terminals of apical dendrites, and sections near the cortical surface show many small unmyelinated, but very few myelinated fibers. In white matter below cortex the numbers of fibers present above one mu decrease up to 3 or 4 mu diameter, while from here to 10 mu the plot of fiber number vs. size is relatively flat, with a decreasing number above 10 mu. All regions of white matter so far studied (cat visual, suprasylvian association, corpus callosum and human frontal) show closely similar distributions (Figure 5). In these regions of cortex it has not been possible to separate afferent from efferent fibers, but in corpus callosum where fibers pass from cortex to cortex in both directions this differentiation is not significant. The distribution here is so similar to that in other subcortical white matter as to suggest that a relatively uniform distribution occurs throughout neocortex. If the fibers above 4 mu are subtracted from the neocortical distribution of cat and human so far studied, the remainder, with certain qualifications, is somewhat similar to the turtle cortex. While the olfactory and other paleocortex has not yet been studied, observations of the hippocampus indicate a paucity of large fibers, and the addition of fibers above about 4 mu is inferred to be a characteristic of neocortex, at least in the higher mammals so far studied.

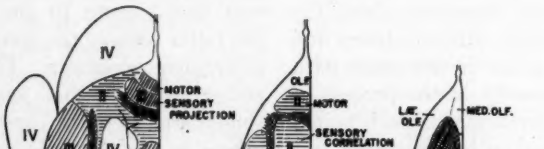


FIG. 5. Stages in the phylogeny of the cerebral cortex. The right diagram represents a primitive cortex, mostly olfactory, the shaded area being the general or somatic cortex. Middle diagram, areas II represent differentiation within the general cortex of areas of sensory correlation; I represents the remainder of the general cortex as association area. Left diagram, III, sensory projection areas of the mammal receive connections through newly differentiated ventral nucleus and medial and lateral geniculate nuclei. IV, expansion in the primate of frontal, temporal, parietal and occipital lobes as elaborations of the association mechanisms. Reprinted by permission from Fig. 5, Bishop, #20, in *Reticular Formation of the Brain*, Little, Brown & Co., Boston, 1958.

Several considerations circumstantially support a view that the larger fibers acquired by neocortex are in fact comparable to those of the more peripheral systems. The optic nerve and tract of the cat (Bishop and Clare, 1955) which is a central tract, shows a fiber size distribution similar to that of cortex, the largest being 12 mu. Fibers from 6 to 12 mu relay to cortex, the smaller fibers to more primitive structures. Again, fibers below 1 mu are abundant. In some lower mammals the optic nerve contains unmyelinated fibers and a process of progressive myelination of even the small central fibers may be inferred.

By comparison, in peripheral nerves of the cat, where cutaneous branches contain no alpha fibers, the largest fibers are 15 to 17 mu in diameter. Afferents larger than this in deep nerves supply the neocerebellar path. Fibers of peripheral nerves decrease in diameter as they enter the dorsal column, which may justify relating the 17 mu maximum of the saphenous to the 14 mu maximum of cortex, and also relating the 6 or 7 mu diameters characterizing the largest (gamma) contribution to lateral column innervation, in the brainstem relayed path (8) to the abrupt decrease in numbers in the cortical distribution at about 4 or 5 mu. It may then be inferred tentatively that the smallest myelinated fibers of cat and human cortex correspond to unmyelinated fibers of peripheral nerves and lower cortices, and the larger fibers acquired by neocortex correspond to a recent addition of fibers of approximately the beta and gamma range of peripheral nerves. The Betz cell area of motor cortex has not yet been examined for fiber size distribution.

The invasion of more recently evolved centers by larger-fiber systems from the periphery

If this process is looked upon as an invasion from the periphery, it might be expected that the projection areas of cortex, so immediately connected with the afferent nerve supply, would show the most numerous complement of larger fibers. This is not the case; all areas of subcortical white matter so far studied, including the corpus callosum, show a closely similar distribution, and even the frontal lobe of the human cortex which might be considered a more recently developed region fails to show an exception. Instead of a literal invasion from the periphery, therefore, the cortex may be considered to have gone through stages of successive development of fibers of increasing sizes *similar to* the developments in the periphery. As a recently developed organ in vertebrates, it appears to have reached its present stage of fiber size development at a later time than did the phylogenetically older peripheral system. It may be significant in this regard that the cerebellum, an older structure than the cerebral cortex, is supplied by postsynaptic projections of the largest fibers from the periphery (alpha fibers in peripheral nerves). Fibers of this size are still missing at the cerebral level, unless indeed the Betz cell fibers from area 4 represent an efferent complementary component.

From the phylogenetic history it might have been anticipated that the thalamus of mammals would contain at least two receptive organs for two somaesthetic sensory paths, one receiving the paleosensory tract which was the only one reaching the thalamus in the premammals, a second receiving the mammalian neo-sensory tract for projection to neocortex. Instead, until recently it seems to have been inferred without much specific knowledge that the newer tracts simply joined the older, especially since part of the classical spinothalamic tract most prominent in man does indeed "join" the medial lemniscus to enter the ventral nucleus. This review of the present information on the courses and components of the sensory systems calls for a re-evaluation of the concepts of sensory modalities and their fiber size correlates.

The vertebrates had large fiber peripheral

afferent detectors before they had acquired a cortex to make elaborate use of them; they added to the primitive cortex a large-fiber system at the same time, perhaps as another phase of a common evolutionary advance, as they extended these peripheral systems to the cortex. In some regions the mammals have relegated to a subcortical functional system some of the afferent machinery available (*e.g.*, the colliculus small-fiber supply from the retina); in other regions they have retained them as components of a more inclusive sensory pattern as in the somaesthetic system. It is significant in this respect that the so-called projection areas, those on which the body image or the surroundings sampled at the periphery are first spread out with a degree of specificity not apparent with respect to functions in the rest of the cortex, should receive their principal and perhaps only direct sensory afferents from only one system of sensory mechanisms. This should afford a clue to what functions neocortex serves; or better, what new functions and functional apparatus have been added to the old premammalian apparatus. The emphasis appears to be on mensuration, intensity, reference to surrounding space, etc.

It cannot be claimed that the dorsal column sensory input dominates the sensory component of cortical behavior, however directly or exclusively this input enters by way of the specific sensory areas. In terms of sensation, nothing is more potent and dominating than pain. It cannot be said that other than dorsal column senses fail to arouse in consciousness a fairly accurate measure of quality, intensity, localization and other parameters of sense. The mammal, and the human particularly, surpasses in this ability the lower animals; that is, his analytic capacity is greater. This may correspond to the fact that his cortex not only contains more neurones, but that he has acquired everywhere, not only in the specific sensory areas, a large fiber component that may contribute its capacities to the functioning of all parts of all material. Insofar as other than the immediately projected senses reach the cortex, or the complex of cortex and thalamus, they should become available to such mechanisms for improved analytical function as may have been associated with the acquisition of the large fiber system.

Whatever scheme of functioning of cortical large and small fiber systems corresponds in general to the facts of central organization, cannot be determined in detail at present. Further clues to the significance of cortical organization in higher animals may be provided by the phylogenetic history of cortex and of the afferent systems associated with it.

STAGES IN THE EVOLUTION OF THE MAMMALIAN CORTEX

This account follows in general Herrick's interpretation (15, pp. 91–108 and 17, pp. 368–383).

Premammalian vertebrates exist which have no true cortex, and in which the anterior pole of the nervous system consists essentially of an olfactory sensory aparatus. As such it makes connections with the diencephalon (chiefly hypothalamus and epithalamus) and with the brainstem integrating apparatus for motor activity, as do other afferent paths (for instance, the visual as noted above). Within this olfactory sensory system a precursor of cortex developed as a facilitator or elaborator of olfactory bulb (sensory nucleus) function (Figure 5). Of this primitive pallium, one area became secondarily connected with the thalamus. This area presumably furnished to the thalamus, already a sensory integrating center, a facilitatory apparatus such as the olfactory cortex had previously supplied to the olfactory bulb. This region of an otherwise olfactory pallium is termed the general cortex.

The next stage in cortical evolution involved a certain degree of specialization within the general cortex. Certain areas of

this region became more specifically related to areas of the thalamus which in turn received impulses from the main sensory systems, visual, auditory and somaesthetic. The olfactory input still entered olfactory cortex, now the specific precursor of the pyriform area and uncus region of man. It still so enters in higher animals and thus reaches cortex primarily by way of the pyriform lobe rather than indirectly by way of the thalamus as do the rest of the sensory systems. Meantime there was established the intimate coordination between the functioning of somatic cortex and dorsal thalamus, which has continued throughout the development of the land vertebrates.

A further advance was accomplished in the development of the mammalian neocortex from its precursor. This consisted, among other differentiations, of the development within or in relation to the sensory regions noted above, of the specific projection areas of visual, auditory and somaesthetic cortex. A corresponding development in the thalamus consisted in the specialization and partial isolation from the rest of the dorsal thalamus, as sensory relays to cortex, of the dorsal nucleus of the lateral geniculate body, of the medial geniculate, and of the posteroventral nucleus. Corresponding differentiation of cerebellar afferent connections to, and of motor outflow from the cortex occurred parallel to this sensory specialization, but do not concern us here. This process might have been expected to leave other regions of thalamus and cortex with relations still resembling those of the premammals, and there are indications in the functional relations of cortex that this has in fact occurred. The surrounds of the projection areas, such as visual para- and peristriate, might be considered as persisting remainders, however further elaborated, of the premammalian regions of sensory reference. Other thalamic nuclei and the association cortex with which they make connections may then be viewed as related to the more generalized somatic cortex, and in this sense to be more primitive in origin and relations than the rest of the thalamocortical apparatus.

As indicated previously, the extension of the various fiber-size paths centrally has occurred progressively, and the larger fiber sensory paths project through the newer areas (more recently specialized) of thalamus to cortex. If the medial spinothalamic path still present in mammals is in fact the persisting paleospinothalamic of the lower vertebrates, its termination in intralaminar nuclei or elsewhere suggests that these constitute an older region of thalamus, and this notion of phylogenetic age might be extended to the layers of cortex with which these nuclei are only indirectly connected. The extension of the peripheral large-fiber system through the medial lemniscus, ventral nucleus and projection area of cortex and its correlates in other special sensory systems of visual and auditory function, then parallels the acquisition by neocortex everywhere of the large-fiber system of the peripheral nerves. Thus two lines of evidence, the extension to thalamus and cortex of sensory paths which consist of progressively larger fibers, and the changes which produced the neocortex itself, support the identification of the successive stages in the evolution of three overlying and reduplicated layers of the sensory apparatus of the mammal.

Cephalization; the relation of the central nervous system development to the specialization of the main sensory systems

It must be granted that a large-fiber overlay in central structures is not the only process by which cortex, for instance, has progressively elaborated its functioning. But it is a general principle of vertebrate system evolution (15) that successively higher level structures have been added in relation to the successive development and elaboration of more cephalad special sensory levels; the cerebellum for instance derives from 5th and 8th nerve nuclei, the

thalamus from the 2nd nerve apparatus, the cortex from an olfactory sensory precursor. What the preceding analysis proposes is that within each sensory level, and extending into the structures derived from these, an increase of fiber size has followed up the neuraxis as a part of the same process of cephalization initiated by, or correlated with, the successive elaboration of centers related to the special sensory systems.

The significance of fiber size; integration vs. analysis

In the frame of reference of sensory system evolution, but with particular reference to cortex, several questions may be raised. What is the significance of the parallel phenomena of the extension of large-fiber paths through the thalamus, the previous sensory target, to cortex, and the acquisition everywhere in cortex, and not only in the sensory projection areas, of an overlay of large-fiber components? And how can we interpret the involvement of only certain "projection" areas of cortex in the particular task of receiving and manipulating not all sensory inputs, but primarily the input of one new path from the periphery?

The large fiber systems, present at the cord level long before the development of cortex, must have contributed something of advantage to *spinal* activity. Perhaps one can surmise, from the results of investigation of spinal activity, what this advantage was in general. The reflexes activated over unmyelinated fibers are diffuse and nonspecific, as the sensations mediated over the same fibers are poorly localized and evaluated; those activated over larger fibers are relatively specific and limited, as the corresponding senses are precisely and accurately localized and quantitatively estimated. One is reminded of Coghill's concept of total-pattern complex behavior *vs.* partial pattern specific reflex response; his point was that complex behavior was not merely a summation of reflex unit activities. Rather, reflex units, if not arbitrary constructs of experimentation, were partial patterns involving limited segments or centers of a nervous apparatus which was in general integrated into activity more complex than reflex level behavior. The total pattern was the primitive one; the reflex was a specialization that could be isolated from the larger patterns by limited and local stimulation.

Both in the sense that small-fibered systems seem to be more primitive than large and to the extent that the more specific, limited and immediate reflex responses tend to follow the experimental stimulation of larger fibers (consider the knee jerk) a correlation may be proposed extending Coghill's concept. This is that the more specific limited unit reflex responses may be not merely more or less arbitrarily isolated fractions of essentially more complex behavior. They may also involve an overlay of new paths more specifically connected, superposed upon a more pervasive system devoted to more generalized types of reaction. The extension of a similar overlay to the higher control systems should then correspondingly increase the detail and specificity of the peripheral efferent responses on which the higher centers operate. Extended to the cortex, an organ initially designed to facilitate and coordinate lower activity in what Herrick has denoted an integrating function, the large fiber system, projecting to the specific sensory areas newly developed in cortex seems appropriate to contribute to its analytic function. In fact the large-fiber dorsal-column afferent system mediates those impulses which the cortex employs particularly for fine discrimination, acute perception and quantitative measurement. What the cortex does by way of analysis of the impulses from the larger peripheral fibers is analogous to what the reflex level does with them; it responds in detail, for particular rather than generalized reaction.

These proposals inevitably raise the question as to just what potentialities are pres-

ent in a large fiber as such, that could enable it (and the neurone which generates it) to offer such advantages as have been attributed to it. The larger the fiber, the higher is its speed of conduction, and it has been proposed repeatedly that shortening of reaction time must be one of the advantages afforded. That the nervous system makes use of large fibers for this purpose seems almost obvious, but this is not necessarily the only property taken advantage of. Another one that may be proposed is power of response, correlated with the amount of current generated by the impulse. This would be of advantage in promoting specificity of response, where one or a few neurones activate others across synapses in a one-to-one sequence, as in many relay and reflex patterns. In contrast to this is the situation, such as that at the terminal dendrites of cortical pyramid cells, where one volley of any strength fails to fire the neurones immediately, but only facilitates their excitation via other paths. These dendrites are activated by many small fibers, and the requirement of many summated impulses to activate the neurone to discharge is appropriate to an integrative action. The largest fibers in the cortical afferent systems on the contrary activate neurones at or near their cell bodies, which then respond immediately by sending all-or-none impulses over their axons. This proposal implies that the synaptic activity at even the fine terminals of large fibers is more potent than that at the terminals of small fibers, a condition concerning which we have insufficient evidence to judge.

THALAMOCORTICAL INTERACTION

Granting that the larger-fiber afferent systems are those most directly innervating the cortical projection areas, it is usually presumed that sensory information mediated over other paths, apparently terminating at lower levels, is somehow represented at the cortical level. How and where do these various paths, mediating similar but qualitatively distinguishable varieties of information as to touch, pain and temperature finally merge at the substrate of consciousness? It is obvious that animals without cortex show behavioral signs of something, if not corresponding to consciousness, at least corresponding to differentiation of sensory modalities by their appropriate reactions. There may have been a stage in evolution when this sort of function was performed, and perhaps limited to as low a level as medulla and brainstem, where much complex behavior is still integrated and organized on a level below consciousness in higher animals. This function must have been more or less transferred to and represented in the thalamus (as the dominant level of integration), when the dorsal thalamus developed as an afferent integrating center out of a diencephalic sensory nucleus. As has been noted above, with the evolution of a somatic cortex from the olfactory system this cortex and the dorsal thalamus have developed in parallel and with profuse interconnections. One must ask, not only how the highest functions of the lower animals were transferred to cortex, but what could the addition of cortex contribute to what the thalamus was already accomplishing. Since it is not possible to answer this question at present in terms of physiology, one may at least follow the physiology of the afferent paths to cortex in search of a possible clue.

The course of the C fiber path to thalamus or cortex is not known, but reticulothalamic paths are known which might serve to extend it upward (25). At least this path would not be expected to be extended through the ventral nucleus. As a path of primitive origin it may be inferred to terminate more like its successor in the premammals, the paleospinothalamic tract. As has been noted, the paleospinothalamic tract terminates in a region of medial thalamus which, when itself stimulated, induces recruiting waves in cortex. Stimulation of many other regions of thalamus also ini-

tiates similar responses, as does stimulation of cortico-cortical paths, although the record is not always so simple and categorical as it may be following medial thalamic stimulation. The activity represented in the recruiting responses appears to be the most primitive of the several patterns recognizable in cortex and is the most prominent of those patterns recorded from the cortex of the turtle.

Primitive (integrative) and more recently elaborated (analytic) functions of periphery and of cortex

Recent examination of these recruiting responses (Figure 6, and Clare and Bishop) indicates that they are characterized by presynaptic fibers from thalamus which are very slow in conduction, about 1 meter per second or less, and which terminate on apical dendrites of pyramid cells near the cortical surface. A second type of response which does not recruit has a shorter latency indicating afferent fibers of higher conduction velocity and corresponding larger diameter. These two types of response and their variants are present in the turtle brain (Orrego, current work) and here the smallest fibers with slowest conduction rate are unmyelinated. In the cat two other types of response are present not found in the turtle, both of still shorter latency. One of these is observed in the projection cortex of mammals and consists of the specific sensory response initiated from the periphery via relays in ventral nucleus, etc. The other is aroused in association areas by stimulation in projection areas.

The recruiting responses and the forms of response allied to them are by various criteria primitive forms of interaction between thalamus and cortex, which persist in higher animals. A regions of thalamus from which *only* these types of response are initiated (the medial nuclei dealt with by Morison and Dempsey) corresponds in this respect to a primitive center. The responses at the cortical level occupy the superficial region of cortex, and occur in the terminal regions of apical dendrites of the pyramid cells. This is in contrast to the large-fiber specific sensory responses of the projection areas which are initiated at the cell bodies of pyramid cells in the IVth layer of cortex. The contact of fibers from a distance with apical terminations of pyramid cell den-

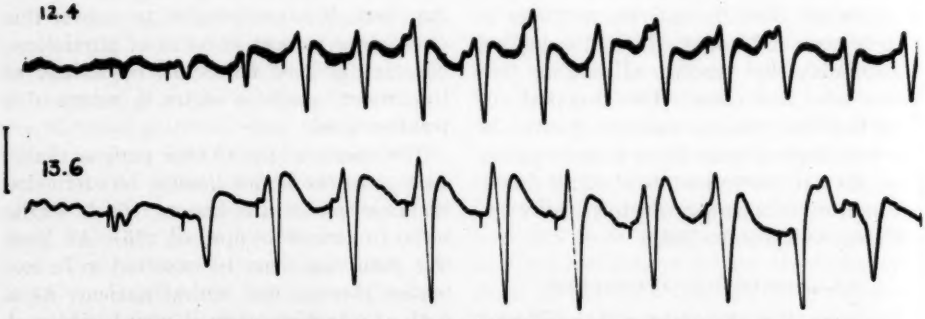


FIG. 6. Responses of cortex to thalamic stimulation at 6 per second. The lower record resulted from stimulation slightly deeper in the thalamus than that for the upper. Both records from the same position of the somaesthetic area, stimulation near the boundary of the dorsomedial nucleus of the cat. Following the first shock artifact two downward-directed, surface-negative waves appear. The earlier decreases in amplitude with repetition of the stimulus, the second increases (recruits). The shorter latency of the first wave indicates responses to stimulation of larger more rapidly conducting fibers. The same two waves appear in the turtle cortex, and in the cat cortex when other areas of cortex or thalamus are stimulated. Stimulation of the sensory radiations of the thalamus give a quite different record (Fig. 4, right column) and the radiation fibers stimulated are larger. Reprinted by permission from Fig. 1, Clare and Bishop, Potential waves in cortex, E. E. G. Clin. Neurophysiol., **8**: 583–602, 1956.

drites is a prominent feature of the primitive cortices of Amblystoma, Necturus and their allies (15). Recruiting type responses from stimulation of the medial thalamus are distributed over a considerable area of cortex. It would appear that any sensory message, if transmitted to cortex from the medial paleothalamic tract, must be transmitted by such paths, and that what this region of the thalamus contributes to sensory perception must be closely associated with this small-fiber apparatus connecting it with cortex.

By contrast the neothalamic relay path from the dorsal column to cortex involves the most recent and complexly organized response of neocortex. It may be significant in this connection that when vision developed in the mammalian cortex, conscious perception of visual activity became so exclusively dependent on this new path that the human after removal of visual cortex is blind. This indicates that if more primitive "visual" paths are involved in mammalian perception at all, they have been relegated to an auxiliary function. Removal of somaesthetic cortex, on the contrary, does not abolish pain and temperature perception, *i.e.*, lateral column senses mediated via smaller fibers. We thus find by comparison of these two otherwise anatomically homologous systems—somaesthetic and visual—different degrees of influence of the older sensory components on the structures responsible for conscious perception. It may be significant in this connection that while a pathway is known from cortex to superior colliculus in mammals, no path appears to course in the reverse direction. Some of the auditory afferents on the contrary relay in the inferior colliculus as well as in the medial geniculate nucleus.

It is improbable however, that the visual cortex is alone the seat of perception; it is surrounded by cortex chiefly involved in further elaboration of visual function, para- and peristriate in man and corresponding association areas in the cat. It seems to be,

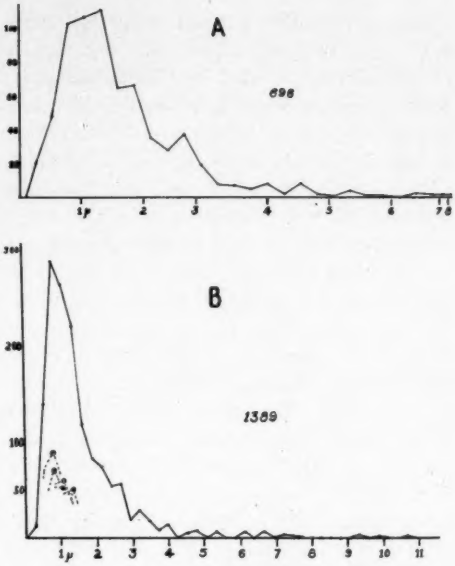


FIG. 7. A, fiber spectrum distribution in the white matter immediately below the human frontal area of cortex, and B, below the cat visual cortex. The dotted curves show the maxima for 5 small areas, added to make the full-line curve. These and other areas so far studied show a maximum number of fibers at just below 1 mu in diameter, a minimum of 0.3 mu and a maximum of 12 to 14 mu, with only a few fibers above 8 mu. The numbers of fibers decrease rapidly from 1 toward 3 or 4 mu, above which the distribution flattens. The turtle cortex lacks these larger fibers of the mammal.

phylogenetically considered, an extra link, a long-circuit around an older thalamic connection. This relation is more obvious in the somaesthetic system. Here the ventral nucleus projection to cortex adds a path for certain discriminating senses to those older paths which are still functional as mediators of characteristic sensory qualities. The somaesthetic projection area interposes an extra segment of pathway between the thalamus and the general association areas of cortex, contributing a preliminary functional analysis of the particular afferent information transmitted over the medial lemniscus. The overall result is to afford a finer discrimination with respect to certain aspects of sensation than is accomplished by

the more primitive pathways through other regions of thalamus. The same function is presumably performed by visual and auditory projection areas. Here certain discriminatory aspects of function are also added in parallel to the more primitive versions of sensory information, and in these two systems, of vision and hearing, the more primitive versions seem to lack any significant cortical representation at all. However, this would be an extreme point of view, without further exploration of the adequacy of cortical vision after destruction of other "visual" paths to thalamus, brain stem, etc., or of hearing after total removal of all auditory cortex (9). While these other paths may not be capable of mediating perception of a modality alone, it would probably be found that they contributed to the final quality or effectiveness both of the sensory effect and of the recipient's reaction to it. If the brainstem as an "activation" center is more than, or different from, another sensory path, it may act as a link in a nonsensory afferent path parallel to those paths that lead more directly to cortex via the projections, as an extension from a stage of evolution when no cortex was present to utilize information on the state of activity in the lower centers.

Central convergence of sensory paths; the thalamocortex as sensorium

Whatever may be the rationale of the sensory contributions of those paths projecting to medial thalamus, the paths projecting more immediately to cortex must finally extend beyond the cortical projection areas to convergence with them. This obvious extension is through the association areas of cortex; the extensions of the older sensory projections to thalamus might also reach association cortex via the recruiting type of path, especially since the recruiting responses are prominent in the association areas. This may not demand, however, that sensation be finally registered, that is, that perception occurs, only in cortex, but merely that two types of paths reach a common corticothalamic complex from two directions, an older one via the "older" thalamic nuclei and a more recent one via the cortical projection areas. What particular regions or networks of neurones might be involved as the ultimate substrate of consciousness is not specified by such a proposal, but only that this thalamocortical complex seems a probable region of final convergence for afferent paths of different modalities, anatomical courses and physiological properties. What complex of facilitations, re-enforcers or other types of auxillary processes may be represented in the rest of the cortex is beyond the scope of this discussion, but further investigation from this point of view would be an appropriate sequel to it.

SUMMARY

The shift in point of view represented in this discussion is from the notion of one fiber size group mediating one modality or quality of sensory effect, (to be sure, with some discrepancies) to a scheme of phylogenetic overlays inherited from our ancestors, each successive stage of which contributed a multiple sensory supply adequate to its possessor. It is, of course, not possible to infer from the sensory effect of activity in these systems as they finally culminated in primates and man, just what functions or at least what sensations were mediated by them in the lower animals where they were first developed. In two of these sensory layers, an unmyelinated component and a small myelinated component, the information seems to have been in effect duplicated, the second version presumably with some refinements not possessed by its precursor. The third layer, of largest sized fibers, is still not complete in mammals not only in the sense that not all modalities are represented, but also in the sense that the largest sizes of fibers project as yet no higher than the cerebellum. Rather, this afferent component has specialized in a particular type

of refinement of only certain afferent functions. In each stage of advance in sensory development, a new center was elaborated to receive and process the new information conveyed over new and presumably more effective afferents, and in each stage the connection to the higher center included an extension of paths already functional at lower levels of evolution, as well as the addition of a new and more direct path. The increase of fiber size seems to have progressed from the periphery forward, as new centers developed in succession rostrally—cerebellum, tectum, thalamus and cortex—in relation to the important sensory systems of spinal and fifth nerve, optic and olfactory levels. The peripheral nerves, for instance, of the frog and turtle contain approximately the same fiber sizes, distributed into the same size groups, as do those of the higher animals, including man. In the turtle a cortex considerably more sophisticated than the frog's still contains only small fibers whose stimulation initiates only the simpler of the activity patterns found in the mammalian cortex. The latter still receives no information directly from the largest fiber component of peripheral nerves, which component has, however, reached the neocerebellum as the dorsal spinocerebellar tract. The cortex even of the primate is still the youngest level of the central nervous system, as the adolescent behavior of even mature humanity may attest. Although it may take time, there is still a prospect for further evolution.

The cortex of the mammal has accomplished a dual advance over its predecessors, furnishing both a facilitator or modifier of thalamic functions and finally in its projection areas it has developed a receptor or direct target organ for a new and improved sensory supply. It is still open to question whether or in what respect both thalamus and cortex are involved in ultimate perception, of any or all the information supplied by the various sensory inputs, but the courses of these various sensory paths suggest their convergence in a structure consisting of both thalamus and cortex, functioning perhaps erratically but inextricably as one organ.

The finding of several modalities of sensory path within a common group of fibers appears to compromise the idea of specific fiber sizes for each sensory quality, or modality if one still permits himself to employ a term which has already become somewhat disreputable from long abuse. It has never been claimed in fact that all the fibers of a given size group served the same identical modality, but rather that all the fibers serving one modality were found within certain size limits. The finding of two pains with different fibers group relations was premonitory; the finding of further dual representation and of a third representation of touch does compromise either the definition of modality or the assignment of any one modality to a given size range. These, however, are largely semantic difficulties, and subjective sensations are even more difficult to assess than are the objective paths related to them. Suffice it for the present purpose however to trace three layers of sensory afferents through their principal central courses and to infer from their structure and physiological properties and behavior what one can as to how the nervous system of the mammal has been put together. If the argument becomes to some degree metaphysical, it is because the action of the nervous system transcends our present physical knowledge of it, whatever the ultimate explanation of its functioning may turn out to be.

"O human head! O magic box!
O wondrous can from labels free.
If what man wants is fame or rocks,
He'll get them when he uses thee."
(*Walt Mason.*)

REFERENCES

1. Bickford, R. G. Experiments relating to the itch sensation, its peripheral mechanism and central pathways. Clin. Sc., **3**: 377–386, 1938.
2. Bishop, G. H. The skin as an organ of senses

with special reference to the itching sensation. J. Invest. Dermat., **11**: 143–154, 1948.

3. Bishop, G. H. and Clare, M. H. Organization and distribution of fibers in the optic nerve of the cat. J. Comp. Neurol., **103**: 269–304, 1955.
4. Bishop, G. H. and Landau, W. Evidence for a double peripheral pathway for pain. Science, **128**: 712–714, 1958.
5. Bowsher, D. Termination of the central pain pathway in man; the conscious appreciation of pain. Brain, **80**: 606–622, 1957.
6. Bruesch, S. R. and Arey, L. B. The number of myelinated and unmyelinated fibers in the optic nerve of vertebrates. J. Comp. Neurol., **77**: 631–665, 1942.
7. Clare, M. H. and Bishop, G. H. Potential mechanisms in cat cortex. Electroencephalog. & Clin. Neurophysiol., **8**: 583–602, 1956.
8. Collins, W. F. and O'Leary, J. L. Study of a somatic evoked response in midbrain reticular substance. Electroencephalog. & Clin. Neurophysiol., **6**: 619–628, 1954.
9. Diamond, I. T. and Neff, W. D. Ablation of temporal cortex and discrimination of auditory patterns. J. Neurophysiol., **20**: 300–315, 1957.
10. Douglas, W. W. and Ritchie, J. M. Nonmedullated fibers in the saphenous nerve signalizing touch. J. Physiol., **139**: 383–399, 1957.
11. Gasser, H. S. and Erlanger, J. The role of fiber size in the establishment of a nerve block by pressure or cocaine. Am. J. Physiol., **88**: 581–591, 1929.
12. Graham, D. T., Goodell, H. and Wolff, H. G. Studies on pain; the relation between cutaneous vasodilation, pain threshold and spontaneous itching and pain. Am. J. M. Sc., **234**: 420–430, 1957.
13. Hall, K. R. L. Pain and suffering. South African M. J., **31**: 1227–1231, 1957.
14. Heinbecker, P., Bishop, G. H. and O'Leary, J. L. Analysis of sensation in terms of the nerve impulse. Arch. Neurol. & Psychiat., **31**: 34–53, 1934.
15. Herrick, C. J. *The Brain of the Tiger Salamander*. Univ. of Chicago Press, Chicago, 1948.
16. Herrick, C. J. Optic and post-optic systems in the brain of Amblystoma tigrinum. J. Comp. Neurol., **77**: 191–353, 1942.
17. Herrick, C. J. *The Evolution of Human Nature*. Univ. of Texas Press, Austin, Tex., 1956.
18. Herrick, C. J. and Bishop, G. H. A comparative survey of the spinal lemniscus system. In *Reticular Formation of the Brain*. Jasper, H. H., ed. Little, Brown and Co., Boston, 1958.
19. Landau, W. and Bishop, G. H. Pain from dermal, periosteal, and fascial endings and from inflammation. A. M. A. Arch. Neurol. & Psychiat., **69**: 490–504, 1953.
20. Lloyd, D. P. C. and McIntyre, A. K. Dorsal column conduction of Group I muscle afferent impulses and their relay through Clark's column. J. Neurophysiol., **13**: 39–54, 1950.
21. Magoun, H. W. and McKinley, W. A. Termination of ascending trigeminal and spinal tracts in thalamus of cat. Am. J. Physiol., **137**: 409–416, 1942.
22. Magoun, H. W. An ascending reticular activating system in the brainstem. A. M. A. Arch. Neurol. & Psychiat., **67**: 145–154, 1952.
23. Mehler, W. R. The mamalian "pain tract" in phylogeny. Anat. Rec., **127**: 332, 1957.
24. Morison, R. S. and Dempsey, E. W. A study of thalamorcortical relations. Am. J. Physiol., **135**: 281–292, 1942.
25. Nauta, W. J. H. and Kuypers, H. G. J. M. Some ascending paths in the brainstem reticular formation. In *Reticular Formation of the Brain*, Jasper, H. H., ed., pp. 3–30. Little, Brown and Co., Boston, 1958.
26. Petren, K. Ueber die Bahnung der Sensitivität im Rückenmarke besonders nach dem Fällen von Stichverletzung studiert. Arch. Psychiat., **47**: 495, 1910.
27. Ranson, S. W. and Billingsley, P. R. The conduction of painful afferent impulses in the spinal nerves. Am. J. Physiol., **40**: 571–584, 1916.
28. Schilder, P. and Stengel, E. Asymbolia for pain. Arch. Neurol. & Psychiat., **25**: 598–600, 1931.
29. Shelley, W. B. and Arthur, R. D. The neurohistology and neurophysiology of the itch sensation in man. A. M. A. Arch. Dermat., **76**: 296–323, 1957.

ADVANCED TRAINING FOR PSYCHIATRIC TEACHING AND RESEARCH

JURGEN RUESCH, M.D.[1]

Although the postwar years have brought about a phenomenal increase in the number of practitioners trained for psychotherapeutic office practice, no parallel development has taken place in the procurement of future teachers, research workers, or personnel qualified for leadership in the field of psychiatry. Inadequate pay, rigidity of organization, and suspicion of new approaches have driven capable and independent young physicians away from institutions and into private practice. This trend of course reduces the probability of significant advances which would materially lower the number of the mentally ill. The conditions which contribute to the scientific stagnation of psychiatry are, among other things, related to the failure of higher education in general (26) and of psychiatry in particular (25). At present, we are confronted with the paradoxical fact that there are more funds available for research than there are takers who are really qualified. If psychiatry is to become professionally and scientifically outstanding, original contributions along theoretical, therapeutic, and scientific lines of endeavor are needed. In order to make this possible, the writer suggests the creation of departments for the advanced training of professionals in the area of psychiatric research, teaching, and therapy. Advanced training ought to have the highest priority on the programs of state and federal government and private organizations (31).

UP TO NOW, ADVANCES IN PSYCHIATRY HAVE BEEN OF A SOCIAL NATURE, CONTRIBUTED BY MEDICALLY TRAINED PSYCHIATRISTS

Since the days of Pinel, the great advances made in the understanding of people and the treatment of mental disease have not originated with engineers, chemists, or pharmacologists. Their technical innovations have not altered the incidence of mental disease and their contributions have proven to be short-lived. Instead, progress was brought about by clinicians who studied the biological, psychological, and social processes that govern man's functioning within himself and with others (25). The only statistically verifiable progress made in psychiatry in the past twenty-five years (with the exception of work in neurosyphilis and pellagra) concerns the duration of hospitalization in the initial phase of the disease. Apparently many of the recommended therapies—electroshock, insulin, and drugs, all of which are non-specific—either necessitate or increase human contact or make the patient amenable to the social approach so that his stay in the hospital is shortened. None of the treatment methods have statistically changed the prognosis of patients who have been hospitalized for one year or longer (16). There is, however, a possibility that the new tranquilizing drugs might change the picture, although it is too early to evaluate the epidemiological effects of the new drugs (4). But therapists have seen cases in which the patients in question resumed their arrested growth when psychotherapy or psychoanalysis was extended over many years. From all this we must conclude that the rehabilitation of the psychiatric patient revolves around the social aspects of his existence and that the methods employed must in the last analysis impinge upon his communicative behavior.

Much of higher psychiatric education is

[1] The Department of Psychiatry, University of California School of Medicine, and The Langley Porter Neuropsychiatric Institute, San Francisco.

subsidized by foundations or governmental agencies, and some of the educational policies are therefore set by consultants from other fields. Some psychologists, anthropologists, sociologists, biologists, physiologists, and anatomists believe that they can "save" psychiatry with their own methods; and successively we have seen a series of fashion trends evolve, each of which in its own right contributed information but did not reduce the numbers of mentally ill. Unfortunately these specialists, who have had an excellent technical training, are rarely familiar with the strategic problems of psychiatry or with the clinical method. Aside from the ataraxic drugs, which were synthesized by various pharmaceutical concerns, all therapeutic methods and psychiatric theories have been developed by medically trained psychiatrists. In view of these facts, the expenditure of money and effort to train psychiatrists at an advanced level seems to be strategically indicated and historically justified.

ADVANCED TRAINING MUST STRENGTHEN THE PSYCHIATRIST TO MANAGE PSYCHOLOGICAL AND SOCIAL CONFLICTS AND TO WITHSTAND SCIENTIFIC DOUBTS

Psychiatrists and psychotherapists walk on a razor's edge between conscious and unconscious, intellect and emotion, animal and machine, sense and nonsense. Unfortunately the provisions available for the management of conflicts of man in the machine age do not seem to apply to certain patients; otherwise they would not seek help. Frustrations incurred because self-realization is barred or the sense of identity is incomplete cannot be overcome merely by raising the standard of living. Our technological and materialistic civilization offers physical comforts rather than spiritual guidance, and the pressure for conformity that modern teamwork requires increases the patient's sense of loneliness. In his work, the psychiatrist cannot avoid facing the fact that many of the patient's conflicts apply to him as well. Inevitably anxiety arises, for which he needs individual supervision. His sense of identity has to be strengthened so that he does not have to join prematurely one of the pressure groups in the field which, through personal support and ready-made rationalizations, assuage the discomfort of their members. In order to advance our knowledge of mental disease, the young man must be able to tolerate threatening conflicts which develop in himself when he begins to question some of the prevailing beliefs. And if he wishes to venture into the unknown, he must be ready to question many of the theories, attitudes, and practices that organized professionalism so vigorously upholds.

Unlike the physicists, who have established a solid cumulative body of knowledge because they agree to what constitutes scientific truth, the behavioral scientists and therapists often deal with problems the truth of which depends upon belief (24). If the physical scientist is in the fortunate position of dealing with few variables at a time and with properties of nature that are not influenced by the attitudes of people, the psychiatrist in contrast is in a position in which human attitudes influence the data. The study of behavior is characterized by the fact that the means of investigation and reporting are the same. Both investigation and reporting necessitate communication which is determined not only by the patient but also by the investigator. Insight into the multiple determinants of human communication usually shakes the very underpinnings of the future professional in that much of what he has taken for granted in the past begins to assume a relativistic meaning. He will discover that without faith in the therapeutic method he cannot help his patients and that if he should be a therapeutic optimist his attitude will distort the truth. Thus he ceases to be an objective scientist. But inasmuch as we live in a scientific era, this conflict hits a tender spot

affecting not only the trainee's sense of identity, his status, and his choice of a subspecialty, but also his pocket book.

INDIVIDUALIZATION OF ADVANCED TRAINING IS PREFERABLE TO STANDARDIZED CURRICULA

The mobilization of creative and leadership talent in psychiatry is a somewhat more difficult task than in other fields. There is no traditional procedure for educating creative people in psychiatry as there is in physics, literature, or physiology. In psychiatry, the young man has to be guided through a world where different philosophies of life meet. He must be protected against the administrator who will flood him with meetings and paper work and who will, through promotions which involve increased administrative duties at an early age, remove him from creative work. He must be protected also against the pure scientist who does not understand the clinical method, much less the problems inherent in living. Also, the young man has to be helped to solve some of his personal problems before he can help patients. Finally, the trainee must be made resistant to the clinician who may be impatient with studies that are not directly related to therapy or who may clamor for immediate results.

The strengthening of the professional identity of the young man is a primary task of advanced education. Only if the young person is capable of resisting the monopolization of his thoughts and activities by various pressure groups will he be able to produce something that is new. To further the growth of the young person, therefore, the method has to be adapted to the trainee's personality, interests, and talents, and the individual should not be squeezed into a standard line of approach. Creativity cannot be mobilized through command performance, and the assignment of professionals to problem areas for which they show no interest has not yielded worthwhile results. In the field of psychiatry, factory type research, mammoth studies, team work, or interdisciplinary approaches so far have not produced basic or fundamental knowledge which would have changed the magnitude of the problem of mental disease. Since in psychiatry the only advances ever made were contributed by well-motivated individuals who pursued their own private interests, our efforts should be directed at avoiding regimentation and standardization.

In postgraduate training, fellows learn mostly from direct contact with senior members. But above and beyond the transmission of information and technical supervision in research, teaching, and therapy, another approach is needed (8). Living in a period in which time-honored traditions are breaking down, and exposed to the multifarious scientific approaches, theories, treatment methods, and opportunities in the field of psychiatry, young men more often than not become thoroughly confused. And if we add to these external difficulties those arising from varying backgrounds, talents, and educational preparation of the young professionals, we come to the conclusion that each trainee has to be helped to resolve his confusion in his own way. Personal guidance by one who is familiar with his work—his therapist notoriously is not—therefore becomes indispensable.

Many of the supervisory hours will be spent in arriving at an understanding of the trainee's difficulties in engaging in creative work. Helping the young man towards a form of creative self-realization is the most important contribution a postgraduate division can make. It is more relevant to the development of research people than training in research methodology or statistics, because if a fellow cannot utilize his resources no amount of formal training will be capable of developing him. It goes without saying that supervision should be carried out by therapists who are thoroughly familiar with the scientific and therapeutic problems of psychiatry and who are

capable of seeing where an individual's personal problems interfere with his work.

But one more qualification is needed. The supervisor has to be a creative person himself who has resolved those conflicts that prevented him from producing. This kind of supervision therefore is not therapy in the narrower sense of the term and does not extend to all the personal problems of the trainee which can be handled separately with a competent therapist. No—this type of supervision takes care of one of the bottlenecks in psychiatric education. It focuses upon the work of the trainee, not from the standpoint of the institution, the research program, or the boss, nor from the standpoint of the patient and the community; this supervision tries to combine some aspects of training on the one hand with self-realization on the other. It is implemented by both individual and group sessions, and the trainee is taught the idea that personality, subject matter, the creative process, and the creative product are inseparably linked.

SUBSIDY OF ADVANCED TRAINING HAS TO REPLACE INHERITED WEALTH

At present, there is an appalling lack of opportunity for young men and women to perfect their skills beyond learning the ABC's of psychiatry (3). If any positions for advanced training are available, they are usually tied to some teaching or service function, with the exception of the few fellowships provided for by government and foundations. And these offer a remuneration which in most cases is totally inadequate. If a young man graduates from medical school at the age of twenty-five and puts in two years of military service, he will have reached the mature age of thirty years by the time he has completed three years of residency training. If he is married and has two children, he would need today—with prices as they are—a salary of $8,129 to maintain the standard of living of a junior professional (12). Should he enter psychoanalytic training, the expense would be nearer to $12,000 a year. Most universities do not pay such compensation at the level of instructor or assistant professor, and the young man turns to psychiatric practice to supplement his income (23). Eventually he may discover that the university center does little to fill his material or intellectual needs, and he begins to identify more closely with the interests of practitioners in the field. At this point, let us remember that other fields of medicine induce physicians to remain connected with teaching centers by furnishing laboratory and hospitalization facilities, quite apart from opportunities for research. But since the practitioner of psychiatry does not need x-rays, operating rooms, special diets, or technicians, there is no reason for him to starve around the training center—particularly because the university often may not even offer him the only thing that he might use: stimulation for his intellectual and therapeutic development. That there exists such a need can be seen from the success that the privately organized psychoanalytic training institutes have had (10, 19).

In the past, the study of medicine was often reserved for those of independent means and therefore salary and length of training were not prime considerations in the young man's career. Now, in contrast, through the advent of scholorships and far-reaching social changes, men and women from all walks of life enter the study of medicine. Since the students marry early and tend to have several children (7), they are hardly able to extend their stay in teaching institutions beyond the absolute minimum that is required. But have the universities taken cognizance of this social revolution? Hardly. The salaries in most places are still geared to the "gentleman student" of yesteryear, whose family provided for him and who had nobody to care for. If we wish to counteract the lowering of standards and if we wish to ade-

quately prepare the young psychiatrist, we have to provide for an adequate income during the training years. In professional education, the gap left by the disappearing professional of independent means has not been filled.

THE CONCEPT OF THE "FULL TIME MAN" IS A LEFT-OVER FROM FEUDAL TIMES

When life, limb, and property belonged to the feudal lord, a serf was truly a full time man. And as long as the social structure required people to work fifty or sixty hours a week, there was little time left for other activities. But in 1958 the five day week predominates and the four day week is in the offing. Also, the scientific schemes of today call for less isolation and more contact with neighboring disciplines. No longer are certain occupations or positions reserved for certain classes of people. In psychiatry in particular, professionals tend to combine administration, teaching, research, and therapy in varying degrees. Together with mass man has come a classless society in which the individual must develop multiple identifications and multiple skills if he wishes to survive. In that sense, nobody is any longer a full time man. In postgraduate training, these social changes have to be acknowledged if education is to be successful.

In accepting the fact that most young psychiatrists enter private practice after they have completed the first three years of basic training, we also acknowledge the fact that we consider their decision to be mature and appropriate. If they were to stay on and neglect the welfare of their wives and children, we would in all likelihood call them irresponsible. But the nation cannot afford to lose talent that has been developed so far. The investment is too great. At present, almost all foundations and government sources tend to refuse half time support for people in private practice. Apparently the men in responsible positions have ignored the fact that practically all knowledge and skill pertaining to the psychological treatment methods have been contributed by such men as Freud, Jung, Adler, and countless others who were in full time private practice. These men have also forgotten that with few exceptions the great majority of advances in clinical medicine have been contributed by European professors, whose activities in the medical schools were and are accompanied by a lucrative private practice.

In an institute for advanced training and research, the place of the clinicians and practitioners is in the teaching of psychotherapy (27, 28). The practitioner contributes to the institution a measure of stability. He is established and stays in one place; he has a vested interest in his community; and he is willing to take responsibilities. Being independent financially, he can afford to evaluate problems without concern for university promotion or civil service rating. The full time system has produced many timid souls and dependent characters who should be counterbalanced by self-employed individuals. In a society in which big corporations, associations, the government, and other institutions control the welfare and often with it the thinking of their members, we need a stock of professionals who do not suffer from the disease of the corporate state (32).

If the government would support practitioners of psychiatry to do part time teaching, the full time men of the training institutions could devote themselves more to scientific investigation. What many experts from other fields do not understand is the fact that psychiatric institutions carry a lot of responsibility for patients and that the staff is heavily taxed in rendering service. What time is left is largely absorbed in teaching the ever increasing numbers of medical students and students from auxiliary disciplines. Therefore, the full time professional has little opportunity to engage in creative pursuits. The reason he could do it in earlier decades was that

there were countless others, working with practically no compensation, who relieved him of routine duties. But this is no longer true.

ADVANCED PSYCHIATRIC TRAINING AND RESEARCH HAS TO BE ORGANIZED IN A SEPARATE UNIT

The educational principles of residency training differ radically from the principles of advanced education. The resident is trained to be a doctor in private practice, state hospital, sanitarium, or clinic. The postgraduate fellow is carefully selected and trained for teaching and research. For him, therapy is not merely a method to be applied but something to be investigated and improved (20). That advanced training can and should be separated from routine training has been proven in the various military staff schools or advanced institutes.

A separate unit for advanced education enables us to undo some of the brain-washing that a youngster has to undergo if he wishes to become a professional. For example, the young man or woman who finishes high school (one in four of the population) and college (one in twenty-five) must undergo further screening. He who becomes a physician is one out of 750 people (30). Thus, through the efforts of colleges, medical schools, and psychiatric boards, the deviants are properly eliminated. After having passed through all the sieves, the young psychiatric graduate usually is eager to learn, well adjusted, but hardly original.

The question may be raised whether the universities are to be blamed for this state of affairs. Teaching institutions are controlled by medical and professional organizations which prescribe qualifications and curricula (5, 6). The profession at large demands conformance to certain stated norms and ideals which are better known as medical ethics; these are rules which prescribe the attitudes to be displayed towards patients and colleagues. Once the young man has taken his M.D. degree and passed the state boards, he is but at the beginning of some five years of further indoctrination. If he wishes to become a specialist, he has to pass the examinations of the specialty board (1), and if he passes the examinations of the American Board of Psychiatry and Neurology he has become one out of 38,000 people (14, 21). If he wishes to become a psychoanalyst he must have a training analysis, attend seminars, be supervised, and fulfill other requirements of the training institute (2). The subject matter the candidate has to concern himself with is circumscribed and his life gets scheduled, including his evenings, leaving little if any time for broadening his horizons and for attending to what one might call "self-realization" (11).

Western civilization does not allot time for meditation, and particularly in an extroverted culture like America introspective pursuits are in disrepute. Once a young psychiatrist has gone through the mill he indeed has become a standard product. But even at this point he does not get a break. If after completion of his specialty training he wishes to make a living and enhance his career, he has to participate in administration, shape psychiatric policies, be a member of committees, organize meetings, and remain in constant contact with the official machinery. But where does that leave the 728,000 patients in psychiatric hospitals and the several millions at large who are looking towards better treatment methods and liberation from disease (22)?

The answer to this question is to train a small segment of the professionals in a different way. If we establish in departments of psychiatry separate units for advanced training, devoted to the education of persons capable of future leadership, an effort has to be made to integrate three of the basic functions of psychiatry: therapy, scientific investigation, and the prevention

of mental disease. In order to develop clinical acumen, therapeutic know-how, and a scientific attitude, the trainee would be largely freed from routine service functions (18). While he is young, the professional should be given the opportunity to spend at least one or two years in the pursuit of knowledge as his colleagues in the departments of physics, chemistry, English, and sociology spend all their lives. The basic idea is that young men and women, if capable, shall be given leisure to think and to create and not be used up prematurely by service or administrative functions. If the physician and psychiatrist is to rise above the level of the vocationally trained technician, if he is to make contributions that are on a par with those in other fields of science, he needs a curriculum similar to that of the well-known research institutions (17). In America, this principle has received recognition in the establishment of the Center for Advanced Study in the Behavioral Sciences in Palo Alto (29), the Institute for Advanced Study in Princeton, the National Institute of Mental Health, Walter Reed Army Institute of Research, or the Mental Health Research Institute in Ann Arbor (9). And the proposal for the establishment of a United Nations Institute of the Human Sciences may still one day be realized (13).

Without wanting to advocate that a well-trained elite is the only way to bring about scientific progress, it is well to remember that outside of research institutions the psychiatrist is hardly exposed to the scientific problems in his field. Most society meetings, lectures, and seminars are geared to the practitioner or the lay person but not to the man who is supposed to advance his field. It has become almost taboo to write or say anything that cannot be understood by the least informed or by the non-specialist. Young professionals therefore are redundantly exposed to the same hash over and over again without ever testing the limits of the known. A postgraduate division for psychiatric training and research would devote its program to pushing knowledge to the borderland of the known and skill to a high level of proficiency.

A division for advanced training and research can benefit all members of an institution. Doctors who hold full or part time service positions could take sabbatical leaves of absence. By undertaking an investigation, by reading and thinking, and by participating in a teaching program they might take back to their jobs a better understanding of the frontiers of psychiatry. And what is a frontier today of course will become routine tomorrow, and thus we would be able to educate a group of leaders whose influence will be felt in all the subspecialties of psychiatry. We could even envisage that a state department of mental hygiene could offer as enticement for meritorious service a six months' stay with such a postgraduate unit. And an administrator might well enjoy leaving his desk behind temporarily and spending six months in an atmosphere of searching scrutiny.

Another distribution of functions is possible if meritorious members of hospital staffs could spend two full days or four half days a week with such an advanced unit throughout the year. This time would be utilized to give them "on the job training," and after a year or two they might be ready to carry on their own investigations. Such time spent at the research center would be used for discussions, seminars, reading, and carrying out a personal research project under supervision.

THE ORGANIZATION OF A DIVISION FOR POSTGRADUATE TRAINING AND RESEARCH

The director of the unit should be responsible for the development and education of the fellows under his tutelage. In such a unit, there would be several full time people, some half time people, and perhaps a few quarter time people. The full time people

would include, aside from the teaching and supervisory staff, the trainees on fellowship grants and doctors on sabbatical leave. The half time and quarter time people would be members who continue with their respective jobs and spend part of the week with the division. Arrangements should be made on an individual basis for each fellow, depending upon the nature of his fellowship, the amount of time available, his personal interests, and his past experience. Applicants should be selected by the director of the division, in consultation with other senior members of the department of psychiatry, who will carefully screen the applicants and admit only those with promise of development. The organization is an informal one. Although the overall curriculum of the fellows is supervised by the director of the postgraduate division, each fellow may be temporarily attached to any service unit or laboratory in the medical school, the neuropsychiatric institute, or clinics and hospitals governed by state, county, city, or private organizations. For that period of time, the fellow is directly responsible to those who are in charge of that particular unit. But notwithstanding this adaptation to the line command (15), the career supervision of the young man rests primarily with the postgraduate division, and particularly with his personal supervisor.

A division for postgraduate training and research could be financed from funds given by the federal or state government or by private foundations. The funds should be granted with the stipulation that each fellow be allowed to devote his time to learning and that he not be required to render service or engage in contract research. With such a method, the department of psychiatry can free its junior staff part time for further learning and insure their professional development. If, for example, a psychiatrist from a mental hygiene unit or a young assistant professor at the university were to draw half-time pay from the division for postgraduate training and remain half-time in his ordinary position, a practicing physician could be hired to fill the half-time teaching job that is left vacant. There are many people in private practice who wish to engage in teaching, and they would be but too delighted to replace an assistant professor who takes a half-time leave for the purpose of engaging in research.

The advantage of developing personnel that is already in a teaching institution lies in the fact that such people have local roots, are committed to an institution, do not have to take their children out of school, do not incur further expenses through traveling, and continue their routine and administrative duties on a reduced scale. They are the persons who will do the teaching in the years ahead, and through their past performance they have already given evidence that they are likely to stay with the institution and that money invested in their future will not be lost.

The director of the division should be a psychiatrist who has shown through his own work that he is capable of creative contributions. No additional positions are required as long as not more than two or three fellows are to be educated. If the number reaches four or five, one additional supervisory position at the associate level becomes necessary. This person should be a researcher in his own right who is capable of inspiring the younger generation. The non-professional personnel should include one top grade administrative assistant who can coordinate schedules, people, and departments as well as attend to the ever-increasing paper work, in order to free other professionals for more productive work. If the number of fellows exceeds four, an additional stenographer should be provided. Research projects which require special supplies or technical assistance will have to be financed from separate funds appropriated by foundations, the department of mental hygiene, or the United States Public Health Service. But it is of para-

mount importance that the hard core of the program—the director, his assistants, and the clerical personnel—be financed from a regular budget which does not derive from special or time-limited funds. Otherwise the continuity of the organization is going to be interrupted.

SUMMARY AND CONCLUSIONS

Although significant progress has been made in the undergraduate teaching of psychiatry to medical students and in the training of residents, the advanced education of professionals leaves much to be desired. To keep young men in and around teaching and research institutions, more effective enticements are needed. Too many men in their fourth to eighth years of specialization are at the present time lost to private practice.

The first effective measure would consist of an increase in salary to compensate the trainee for progressive inflation, taxes, and expenses incurred in psychoanalytic and psychotherapeutic training. Such changes, however, are likely to be made only gradually.

The second effective measure is the creation of divisions for postgraduate training and research, the members of which concern themselves with the advanced education of professionals who have completed their basic training in psychiatry. The advantages of separating advanced from routine training lie in the choice of personnel and the nature of the atmosphere. Only well-motivated people shall be admitted to advanced levels of training and indoctrination shall be minimized.

The curriculum of such a postgraduate division therefore shall not be dictated by associations, specialty boards, or the requirements for higher degrees; but the study plan shall be adapted to the needs and interests of the individual trainee. Personalized supervision in a kind of tutorial system shall be used to mobilize creative talent and to encourage scientific exploration, therapeutic experimentation, and expert formulation.

The practice of separating therapy, teaching, and research into different fields of endeavor is relevant for technical considerations only. Leaders in professional fields are not technicians, vocationally trained, with specialized skills. They are men and women who have to make complex decisions and who therefore require an integrated education. To provide research and teaching personnel, to familiarize service personnel and administrators with the problems of research would be the goal of such a proposed advanced unit for training and research. Only well-trained professionals are likely to guide government, associations, universities, and state hospitals in the establishment of effective therapy, teaching, and research programs.

REFERENCES

1. American Board of Psychiatry and Neurology, Inc. *Information for Applicants; Rules and Regulations of the American Board of Psychiatry and Neurology, Inc.* Rochester, Minnesota, 1955.
2. American Psychoanalytic Association. Standards for the training of physicians in psychoanalysis. Bull. Amer. Psychoanal. Assoc., May, 1950.
3. Bennis, W. G. The effect on academic goods of their market. Amer. J. Sociol., **62**: 28–33, 1956.
4. Brill, H. and Patton, R. E. Analysis of 1955–1956 population fall in New York State Mental Hospitals in first year of large scale use of tranquilizing drugs. Amer. J. Psychiat., **114**: 509–517, 1957.
5. Conference on Psychiatric Education. *Psychiatry and Medical Education.* Report of the 1951 Conference. American Psychiatric Association, Washington, 1952.
6. Conference on Psychiatric Education. *The Psychiatrist; His Training and Development.* Report of the 1952 Conference. American Psychiatric Association, Washington, 1953.
7. Davis, K. The early marriage trend. What's New, No. 207, pp. 2–6, Fall, 1958. Abbott Laboratories, Chicago.
8. Ebaugh, F. G. Graduate teaching of psychiatry through individual supervision. Amer. J. Psychiat., **107**: 274–278, 1950.
9. Gerard, R. W. Problems in the institutionalization of higher education; an analysis based on historical materials. Behavioral Sci., **2**: 134–146, 1957.

10. GRALNICK, A. A postgraduate psychoanalytic training program. Amer. J. Psychiat., **106**: 841–844, 1950.
11. GREGG, A. Narrative for a specialist. Amer. J. Psychiat., **100**: 191–194, 1944.
12. HELLER COMMITTEE FOR RESEARCH IN SOCIAL ECONOMICS, UNIVERSITY OF CALIFORNIA. *Quantity and Cost Budgets for Two Income Levels (Prices for the San Francisco Bay Area, September, 1954)*. Regents of the University of California, Berkeley, 1955.
13. INTERNATIONAL COMMITTEE OF THE SOCIETY FOR THE PSYCHOLOGICAL STUDY OF SOCIAL ISSUES. Proposal for the establishment of a United Nations Institute of the Human Sciences. Human Relat., **1**: 353–372, 1947.
14. JOINT INFORMATION SERVICE OF THE AMERICAN PSYCHIATRIC ASSOCIATION AND THE NATIONAL ASSOCIATION FOR MENTAL HEALTH. *Fact Sheet No. 2*. APA and NAMH, Washington, 1957.
15. JURAN, J. M. Improving the relationship between staff and line: an assist from the anthropologists. Personnel, **32**: 515–524, 1956.
16. KRAMER, M., GOLDSTEIN, H., ISRAEL, R. H. AND JOHNSON, N. A. *A Historical Study of the Disposition of First Admissions to a State Mental Hospital*. Public Health Monograph No. 32. United States Government Printing Office, Washington, 1955.
17. KUBIE, L. S. The problem of maturity in psychiatric research. J. Med. Educ., **28**: 11–27, 1953.
18. KUBIE, L. S. The pros and cons of a new profession: a doctorate in medical psychology. Tex. Rep. Biol. Med., **12**: 692–737, 1954.
19. LOEWALD, H. W. Psychoanalytic curricula—principles and structure. J. Amer. Psychoanal. Assoc., **4**: 149–161, 1956.
20. MARQUIS, D. B. Research planning at the frontiers of science. Amer. Psychologist, **3**: 430–438, 1948.
21. MOHR, P. E., (ed.) *American Medical Directory*. American Medical Association, Chicago, 1956.
22. NATIONAL MENTAL HEALTH COMMITTEE. *What are the Facts about Mental Illness in the United States?* National Mental Health Committee, Washington, 1956.
23. POTTER, H. W., KLEIN, H. R. AND GOODENOUGH, D. R. Problems related to the personal costs of psychiatric and psychoanalytic training. Amer. J. Psychiat., **113**: 1013–1019, 1957.
24. RUESCH, J. AND BATESON, G. *Communication, the Social Matrix of Psychiatry*. Norton, New York, 1951.
25. RUESCH, J. The trouble with psychiatric research. A.M.A. Arch. Neurol. Psychiat., **77**: 93–107, 1957.
26. RUESCH, J. Conformance, originality, and responsibility in higher education. (In press.)
27. SZUREK, S. A. Remarks on training for psychotherapy. Amer. J. Orthopsychiat., **19**: 36–51, 1949.
28. SZUREK, S. A. Teaching and learning of psychoanalytic psychiatry in medical school. Psychoanal. Quart., **26**: 387–396, 1957.
29. TYLER, R. W. Study center for behavioral scientists. Science, **123**: 405–408, 1956.
30. U. S. BUREAU OF THE CENSUS. *Statistical Abstract of the United States: 1956*. U. S. Government Printing Office, Washington, 1956.
31. UNIVERSITY OF CALIFORNIA. Educational policy and research: Report of study committee to All-Faculty Conference. University Bulletin, **6**: 146–150, 1958.
32. WHYTE, W. H., JR. *The Organization Man*. Simon & Schuster, New York, 1956.

EGO STRENGTH AND LENGTH OF RECOVERY FROM INFECTIOUS MONONUCLEOSIS

NORMAN S. GREENFIELD, PH.D.,[1] ROBERT ROESSLER, M.D.[2] AND ARCHER P. CROSLEY, JR., M.D.[2]

The present study, rooted in the view that recovery from illness is an essential adaptive function of the ego, tests the hypothesis that there will be a positive relationship between psychological health and rate of recovery from infectious mononucleosis.

Since the introduction of the precepts of psychosomatic medicine more than three decades ago, the primary emphasis in this field, both in practice and in research, has been concerned with the traditional mind-body relationships in a relatively circumscribed group of illnesses. The classical paradigm has been in terms of the degree to which emotional factors contribute to organic pathology and a great deal of energy has been devoted to the development of theories regarding precise relationships between psychological conflictual areas and organ system changes.

The past decade, especially the past few years, has witnessed the beginnings of significant modifications of the classical psychosomatic approach, especially in the realm of research. The "essential unity of man" is no longer a mere bio-philosophical slogan; concepts like this are being dusted off and subjected to the rigorous scrutiny of scientific methodology. Gradually and painstakingly a new body of knowledge is being accumulated; the rationalism of the theorist is being supplemented by the empiricism of the rigorous investigator. More and more we are on solid footing in the belief that, "psychosomatic" in truth describes two functional aspects of the same structural unity and that indeed all function involves structure in a concordant relationship. As Cobb (4) has pointed out, the traditional separation between "functional" and "organic" disease is an anachronism which has done much to obscure the problems of disease. Strides are being made to take a fresh look at the problem and concepts are being translated into hypothetical constructs which have heuristic purpose.

The notion, at least at the intuitive level, that the course of illness may be related to the state of psychological well-being of the patient may well be as old as man's concern with the welfare of man. But it is only in the recent past that this kind of thinking has been made explicit.

Perhaps the earliest and most extensive work relating psychological variables to the length of recovery from a disease process was with pulmonary tuberculosis. William Osler has been quoted as saying, "It is just as important to know what is in a man's head as what is in his chest if you want to predict the outcome of his pulmonary tuberculosis" (12). The literature in this field is extensive and it would be uneconomical to cite it here in detail. The interested reader may refer to reviews by Daniels and Davidoff (5), Berle (2), and by Day (6).

Perusal of the literature dealing with other diseases suggests that there has been more interest in this line of reasoning than is generally recognized. Schneider *et al.* reported some degree of success in relating preoperative psychological study to recovery from surgical procedures (11). Blumberg, West and Ellis (3) have demonstrated a relationship between personality factors

[1] Department of Psychiatry, University of Wisconsin Medical School, Madison, Wisconsin.

[2] Department of Medicine, University of Pennsylvania Medical School, Philadelphia, Pennsylvania. The authors wish to thank Miss Deborah Katz for her assistance in processing the data for this study.

and the course of neoplastic disease and, dramatically, Klopfer has demonstrated Rorschach differences in patients with slow-growing tumors as opposed to patients with fast-growing tumors (7). Schmale (10) has recently related object loss to onset of somatic disease. In keeping with the presently-held notion that disease and recovery therefrom are facets of a general adaptive function, Lovett Doust has reported that there is a greater incidence of physical disease in the histories of psychiatric patients than in a control group (8).

METHOD

The present investigation tested the hypothesis that there will be a positive relationship between psychological health and length of recovery from disease. Infectious mononucleosis was selected because it follows a fairly predictable course which may be observed through successive hematological examinations and because it runs its natural course free of significant therapeutic amelioration. Additionally and importantly, the disease is fairly common in a university population, resolving some of the problems of gathering an adequate sample with a safe likelihood of availability for follow-up study.

The subjects were thirty-eight undergraduate university students. There were 21 males and 17 females. All subjects were patients of the University of Wisconsin Student Health Service.

The judgment of the duration of the disease was made independently by an internist from the medical records and without contact with the patient. Criteria for objective onset were: 1) a total white blood cell count of more than 10,000 per cubic millimeter, 2) a lymphocyte count greater than 60 per cent, and 3) the presence of atypical lymphocytes. All of these criteria had to be met or, alternatively, that there be a positive heterophile test in the presence of leucopenia. The criteria for the objective termination of the illness were: 1) a total white blood cell count of less than 10,000 per cubic millimeter, 2) a total lymphocyte count of less than 60 per cent, and 3) no atypical lymphocytes.

According to these criteria the mean duration of the illness of the total group was 12.32 days with a standard deviation of 9.23 days. The twenty patients who exceeded the mean recovery time were labeled the "long-recovery group" and the eighteen patients whose recovery time was below the mean were designated the "short-recovery group". Approximately six months after recovery from the disease the Minnesota Multiphasic Personality Inventory was administered to each of the subjects. Two MMPI scales which have demonstrated their validity as measures of general psychological health were employed as operational criteria of ego-strength: the Baron Ego Strength Scale (1) and the Meeker LH-4 Scale (9).

Both of these scales were originally developed to deal with the problems of a psychiatric population. The Baron Scale, consisting of 68 MMPI items, was designed to detect the response of psychoneurotic patients to psychotherapy. In the author's words, "It appears to measure the various aspects of effective personal functioning which are usually subsumed under the term 'ego-strength' " (1, p. 327). Similarly the LH-4 Scale was initially designed to differentiate short-term from long-term hospitalized schizophrenic patients. Here too, the common factor appeared to be one of overall personality strength with short-term patients demonstrating significantly higher scores. The LH-4 Scale consists of 28 items and although it is aimed at the same psychological dimension as the Baron Scale, the independence of the two scales is indicated by the fact that there are no overlapping items.

RESULTS

Both of these independent indices of ego-strength or psychological health differentiated the "long-recovery group" from the

"short-recovery group." The mean Baron ego-strength score for the "long-recovery group" was 48.00 while that for the "short-recovery group" was 50.39. This difference yields a t-score of 1.86 which on a one-tailed test is significant at beyond the .05 level. On the LH-4 Scale the mean for the "long-recovery group" was 19.30 while that for the "short-recovery group" was 21.22 which yields a t-score of 2.13, significant at the .02 level of confidence. Thus, the hypothesis is confirmed at the customarily acceptable levels of confidence.

In order to ensure the functional independence of the two measures for the present sample a product-moment correlation was performed. A correlation of +.47 further demonstrates this independence.

DISCUSSION

Some comments about the methodology employed in this study are pertinent. The differences we have reported reached the usually accepted levels of statistical significance despite the fact that our population was presumably a relatively healthy one psychologically. It seems likely that, had we selected greater extremes in ego strength, these differences would reach even higher levels of significance. In short, it seems to us that we are dealing here with a powerful variable if its influence is measurable under investigative conditions which were considerably less than ideal.

These findings for infectious mononucleosis are in general agreement with those previously cited for other disease processes. Although we are currently extending the methodology of the study to still another disease—viral pneumonia—as a cross-validation, it seems to us that there is now substantial empirical evidence for the assertion that psychological health and somatic health are only arbitrarily separable. To paraphrase a familiar quotation, "Where soma is, there shall ego be".

Assuming the validity of this view, it remains to refine our knowledge of the ways in which soma and ego are systematically related. A preliminary finding of another study soon to be reported would appear to suggest a definite relationship between discriminative psychological perception and differential physiological response. We have reason to believe that the individual who is unable to perceive psychological events accurately will correspondingly be unable to discriminate internal physiological events with accuracy. In other words, a possible explanation for our findings in this study is that those persons with relatively less ego strength cannot discriminatively perceive, and respond to, the various physiological stimuli occasioned by the illness; their adaptive capacity is impaired by inaccurate perception. We are pursuing this latter hypothesis further.

Finally, we wish to emphasize that it is not the intent of this paper to present measures for use in predicting recovery times among infectious mononucleosis patients. Although statistically significant, the differences presented are not substantial enough to be employed practically for individual prediction. Our interest has been in the confirmation of our hypothesis and we believe that the reported differences do confirm it.

SUMMARY

An investigation was carried out to test the hypothesis that recovery from illness is an essential adaptive function of the ego and that there will be a positive relationship between measures of psychological health (ego strength) and rate of recovery from infectious mononucleosis. Using objective data contained in medical records, two groups, one slow in recovering and one recovering rapidly, were defined. Six months after recovery from the illness the Minnesota Multiphasic Personality Inventory was administered to each patient. The two measures of ego strength employed both differentiated the groups at appropriate levels of statistical confidence.

REFERENCES

1. Baron, F. An ego-strength scale which predicts response to psychotherapy. J. Consult. Psychol., **17**: 327–333, 1953.
2. Berle, Beatrice. Emotional factors and tuberculosis: a critical review of the literature. Psychosom. Med., **10**: 366–373, 1948.
3. Blumberg, E. M., West, P. M. and Ellis, F. W. A possible relationship between psychological factors and human cancer. Psychosom. Med., **16**: 277–286, 1954.
4. Cobb, S. *Foundations of Neuropsychiatry*. Williams and Wilkins, Baltimore, 1958.
5. Daniels, G. E. and Davidoff, E. The mental aspects of tuberculosis. Amer. Rev. Tuberc., **62**: 532–537, 1950.
6. Day, G. The psychosomatic approach to pulmonary tuberculosis. Lancet, **260**: 1025–1028, 1951.
7. Klopfer, B. Psychological variables in human cancer. J. Proj. Techniques, **21**: 331–340, 1957.
8. Lovett Doust, J. W. Psychiatric aspects of somatic immunity: differential incidence of physical disease in histories of psychiatric patients. Brit. J. Soc. Med., **6**: 49–67, 1952.
9. Meeker, F. *An Exploratory Study of Specific and General Personality Dimensions Related to Length of Hospitalization Among Psychiatric Patients*. M. A. Thesis, University of California, Berkeley, 1958.
10. Schmale, A. H., Jr. Relation of separation and depression to disease. Psychosom. Med., **20**: 259–277, 1958.
11. Schneider, R. A., Gray, J. S. and Culmer, C. H. Psychologic evaluation of surgical patients. Wisconsin Med. J., **49**: 285–290, 1950
12. Vitale, J. H. *An Investigation of Some Personality Correlates During the Clinical Course of Tuberculosis*. Ph.D. Thesis, Stanford University, 1953.

THE DESIGN AND CONDUCT OF EXPERIMENTS ON THE PSYCHOLOGICAL EFFECTS OF DRUGS

HARVEY NASH, PH.D.[1]

Experiments on the psychological effects of drugs are generally complex. Complexity arises from attempts to limit the many kinds of error to which such experiments are prone. The present paper, beginning with basic facts and principles of statistics, psychology, and pharmacology, goes on to discuss in detail considerations that influence the structure of experiments on the psychological effects of drugs.

BASIC DESIGN CONSIDERATIONS

ELIMINATION OF BIAS

An experiment may begin with such a question as, "Does Compound X relieve pain?" Seeking an answer, the investigator may compare the pain reported by subjects before and after the administration of Compound X. This comparison may of course be biased, for factors other than the drug, as perhaps spontaneous recovery from illness, may have lessened the pain. To reduce the likelihood of bias, results obtained from subjects treated with Compound X are compared with results obtained from "control" subjects not treated with the drug.

However, in clinical studies[2] serving therapeutic as well as scientific functions, the use of control subjects raises obvious problems. As a way of minimizing these problems, Lasagna (44) has suggested that a drug of uncertain potency can be evaluated, in a controlled clinical study, by comparing it with a "standard" drug, whose therapeutic value has been adequately determined. (Morphine has been used as a standard of analgesic potency, for example (42).) By this means, bias can be avoided without recourse to an untreated control group, and therapeutic benefits can meanwhile be conferred by the standard drug, should the new drug prove to be valueless.

In any event, once the investigator has decided to compare a given set of dosage forms, he must assign subjects to the various dosage forms on a *strictly random* basis, if a valid test of significance is to be applied to the experimental findings.[3]

REDUCTION OF EXPERIMENTAL ERROR

A program of research is effective when it yields rich results per unit expenditure of time, effort, or money. A research program to remain effective must continue to uncover fruitful leads without dissipating excessive energy in pursuit of stray leads. To insure himself against too many unproductive side excursions, the investigator rejects results which fail to reach statistical significance at a specified level; the 5 per cent level is selected quite frequently.

[1] Department of Neurology and Psychiatry, Northwestern University Medical School, Chicago. The writer is indebted to the readers of an earlier version of this paper—to K. A. Brownlee, Raymond Weil, and Wayne Dennis, in particular—for their helpful comments.

[2] While this paper is devoted primarily to design problems in studies involving "normal" human subjects, much of the discussion is also applicable to studies involving pathological conditions in human subjects and to studies involving animal subjects. Suggestions which are specifically directed to studies on the therapeutic effects of drugs are contained in several discussions of controlled clinical trials (9, 37, 39, 44, 69, 70) and in the contributions of Beecher and his associates to the study of pathological and non-pathological subjective states (4, 5).

[3] In evaluating the results of a study in which a drug is compared with a placebo or with a standard drug, the *t*-test is applied; when *each* of several drugs is compared with a placebo or with a standard drug, it is most appropriate to apply Dunnett's test (13). When the distribution assumptions for the *t*-test and for Dunnett's test do not apply (10, 15), the data may be transformed (3), or evaluated by means of "non-parametric" methods (54, 65).

Research efforts are also dissipated when drug effects that actually exist are masked by excessive "experimental error." By "experimental error" is meant the random variations in response which are found to occur naturally after medication. Experimental error may be attributed to variations among subjects before the onset of the experiment; to variations—either environmental or internal to the subject—during the course of the experiment; or to measurement error. Thus, variations in reported feelings after drinking a couple of martinis may be due to individual differences in regard to body size or customary alcohol intake, to variations in the emptying time of the stomach, to changes in the social environment, to fluctuations in room temperature or barometric pressure, to changes in composition from one batch of martinis to the next, to spontaneous fluctuations in fatigue or body temperature, or to the unreliability of self-ratings.

It is desirable that the investigator anticipate major sources of error and attempt to control them. He may, for example, establish a controlled laboratory environment for his subjects, regulating room temperature, humidity, illumination, and acoustic noise. Or he may attempt to reduce error by the judicious selection of subjects, by the choice of an appropriate experimental design, or by auxiliary statistical analysis of the experimental findings. Naturally, the effort to control a given source of error should be proportional to the degree of disturbance expected from that source of error. Considered throughout this paper are various ways of minimizing the kinds of error which are likely to disturb investigations of the psychological effects of drugs.

SELECTION OF SUBJECTS

Before subjects can be assigned to treatment conditions, they must first be recruited. This task is less simple than might appear at first glance. If, for example, the results obtained from the experimental sample are to be generalized to a specified population, it is necessary to select subjects in such a way that they truly represent this population. In nonclinical studies, the population of interest is usually "adult human beings in western society," or even "the entire human race." Although it is hardly possible to obtain thoroughly representative samples of such large populations for experimental purposes, fairly representative samples of substrata of these larger populations have been studied occasionally, as in the evaluation of vaccines.

In studying the psychological effects of drugs on human subjects, representative samples are ordinarily obtainable only for "captive" populations, such as those of a hospital or prison. While a so-called captive population may be readily available, and while representative and even complete sampling of this restricted population may be possible, the value of findings obtained from a population of this kind may be quite limited. Thus, results obtained after administration of meprobamate to the entire population of X State Mental Hospital have only limited generality, if the investigator is trying to learn about the anti-anxiety effects of the drug, not so much because the particular hospital may be unrepresentative, but because the anxiety prevalent in institutionalized mental patients may differ markedly, both in kind and in degree, from that typical of patients seen in office practice. The value of findings relating to the psychological effects of drugs may, moreover, be impaired, should the captive subjects prove to be uncooperative, poorly motivated, or irresponsible.

Unfortunately, the main interest of the investigator too often lies in some group other than the one which happens to be available.

In order to avoid the risks associated with captive subjects, many investigators deliberately choose to work with selected volunteer subjects. The data obtained by Lasagna and von Felsinger (44, 47) suggest, however, that an unusually large number of

those who volunteered for certain drug experiments suffered from severe psychological maladjustment, and were therefore quite unrepresentative of the population-at-large. Clearly, there are disadvantages in the use of either volunteer or captive subjects. The choice between these two kinds of subjects must depend on subject availability, on the kinds of psychological tasks employed to measure the drug effects, as well as on the incentives which can be employed to improve the motivation of subjects reluctant to participate.

The responses of the atypical subjects of Lasagna and von Felsinger (47) to the drug treatments appear to have been determined, in part, by the pre-experimental pathology of these subjects. The dependence of drug effects on psychopathology and on pre-experimental differences in personality has been noted by other investigators (6, 38, 49, 68). The effects of drugs are, indeed, modified by a number of characteristics of the individual.

Since differences among individuals affect response to drugs, these differences would also be expected to affect experimental error. Lasagna and von Felsinger (47) did, in fact, note that the responses of their atypical subjects served to increase experimental error (although not to the point of obscuring the typical effects of the drugs under study). The reduction of experimental error, by correcting for personality predispositions or other characteristics of the individual, or by care in the selection of subjects, is discussed later in this paper.

SOME EXPERIMENTAL DESIGNS[4]

The simplest kind of drug study is one in which two or more dosage forms are compared (nitrous oxide-plus-oxygen versus air, for example (66)), under standard conditions. Since drug effects may vary with the circumstances surrounding the experiment, it is often desirable to study the effects of a drug over a range of conditions. A "factorial design" (17), which is an experimental arrangement for studying drug effects over a range of conditions, permits broader generalizations to be drawn with greater efficiency than does a series of separate studies, each limited to a single set of conditions. By furnishing estimates of the interactions between dosage forms and conditions of experimentation, a factorial study also provides information on the conditions which govern the appearance of a given drug effect.

Factorial designs have been used extensively in a series of studies conducted at the School of Aviation Medicine (*e.g.*, 33, 34, 35, 36, 59, 60, 61, 63). In one of these investigations, Payne (59) studied the joint effects of cerebral depressants, motivation, and overlearning on verbal retention. While the main effects of these three factors were not well established in this study, the interactions among the factors did indicate that both overlearning and motivational feedback were able to forestall an impairment of function by the depressants.

The efficiency of experimentation can be further increased by segregating from experimental error the effects of known sources of error which may not be readily eliminated. Dividing the experiment into relatively homogeneous time units, each of which is a miniature version of the experiment as a whole, is one way of minimizing unwanted variations which occur as the experiment progresses. This design (termed "randomized blocks") is quite suitable for controlled clinical studies (with either ambulatory or hospitalized psychiatric patients) to which subjects are admitted in chronological sequence.

Psychopharmacologists have long recognized that observations on the psychological

[4] The investigator with little statistical background could profit from a fuller introduction to the subject of experimental designs, such as can be found in Fisher (17), Finney (16), and Cochran and Cox (11). He would find it useful, moreover, to consult a statistician, not only after his data are obtained, but also during the planning of the experiment, before its design is fixed.

effects of drugs might be affected by the temporal order or sequence of presentation of drugs or of tests. One of the earliest arrangements for eliminating bias due to order of presentation involved the alternation of drug days and non-drug days. Another scheme for balancing order of presentation is the "crossover" (or "switchback") design, in which dosage form A is presented before dosage form B to one half of the subjects, while dosage form B is presented before dosage form A to the other half of the subjects (see Barmack (2); for modified crossover designs, see Goodnow *et al.* (25) and von Felsinger *et al.* (67)).

The crossover design is a special case of the "Latin square" design. The Latin square *resembles* randomized blocks in that it segregates some unwanted variation from experimental error. It *differs from* randomized blocks both in its greater complexity (it segregates the effects not of one source of unwanted variation, but of two independent sources of unwanted variation) and in its greater efficiency. The efficiency of the Latin square is usually gained at the expense of the requirement that all interactions be zero, a requirement which, according to McNemar (51), seriously restricts the use of the Latin square in psychological research. Although there are situations in which interactions are permitted (27), and although small interactions are tolerated, in practice, for the remaining situations, the investigator of the psychological effects of drugs is well-advised to avoid the use of the Latin square unless he has expert statistical guidance. Some of the difficulties attending the use of the Latin square in studying the psychological effects of drugs are treated in a later section on learning.

In the Latin square, all interactions are omitted from the analysis. The efficiency of experimentation can be increased considerably by the deliberate neglect of one or another interaction between dosage forms and conditions of experimentation. Where such interactions are neglected, the effects of these interactions become indistinguishable from the natural variations in the responses being studied, and the neglected interactions are said to be "confounded" with experimental error. For instances of such confounding, the reader is referred to the "split-plot" designs used in several of the studies conducted at the School of Aviation Medicine (33, 34, 35, 36, 61, 63). (The split-plot design, it should be noted, may prove useful for studying temporal variations in drug effects; the interaction of drugs with trials can provide a precise estimate of the effects of drugs on learning, for example.)

The investigator may be interested in comparing different quantities or concentrations of the same drug, to determine the "regression" (or dependence) of drug effects on dosage. Test performance may, for example, be related to drug concentration in the body fluids (22). Drug effects may be related to the quantity of drug administered; thus, Keats, Beecher, and Mosteller (42) indicated that the analgesic potency of a drug can be assayed by relating the dose-response curve to the percentage of pain-relief afforded by a standard dose of morphine. Regression studies are by no means incompatible with the designs discussed above; when different quantities of the same drug are included in a factorial design, for example, it is possible to examine the data for regression of drug effects on quantity of drug (*cf.* 62).

Another technique for increasing the efficiency of experimentation is the analysis of covariance. This procedure corrects experimental observations for the effects of known, but uncontrolled, sources of error (14, 17, 26). The correction is applicable when the response to medication is shown to be dependent on an available accessory measure (such as I.Q.) which is not usually utilized in designing the experiment proper. Although the analysis of covariance is not

performed until after the completion of the experiment proper, the selection of accessory measures requires the same thoughtful planning as is required in the design of the experiment proper. A poor choice of accessory measures absorbs effort better spent obtaining additional responses to medication. Good judgment in the choice of accessory measures is especially important, for the opportunity to collect the desired accessory data may be lost once the experiment proper has begun.

The analysis of covariance is by no means the only procedure for utilizing accessory measures. Accessory measures can also be employed to arrange a matched or balanced design (see Steinberg (66), for an example), or they can be incorporated into an "index of response," such as the difference between the post-drug and pre-drug values of a given variable (41). The "difference-score" method, though somewhat less precise than the analysis of covariance, is in common use in psychological research. (But note that it is only appropriate to subtract measures which are on the same scale.) The use of the difference-score method is usually justified, both because of its computational simplicity and because the underlying assumptions do not usually lead to serious error. See Cox (12) for a thorough discussion of accessory measurements. For instances of the use of the difference-score method in psychopharmacological research, see Goodnow *et al.* (25), Steinberg (66), and von Felsinger *et al.* (67); for instances of the use of covariance analysis to adjust post-drug values for pre-drug values of the same variable, see the studies conducted at the School of Aviation Medicine.

The full potentialities of accessory measurements have yet to be realized in psychopharmacological research, despite the frequent adjustment of post-drug values for pre-drug values of the same variable. Covariance analysis could be used profitably to explore the dependence of drug effects on suggestibility and on other personality traits (31, 49, 68), as well as to correct for any personal characteristics which are found to modify drug effects.

ESSENTIAL DESIGN CONSIDERATIONS VERSUS ECONOMIC DESIGN CONSIDERATIONS

Although some of the basic principles of experimental design have been disseminated widely, one or another of these principles is frequently neglected in studies on the psychological effects of drugs. To be sure, some basic principles are not widely known, while others are difficult to grasp. But even principles which appear to be understood when separately considered may, when interwoven into a design of some complexity, turn out to be a source of difficulty for the investigator.

Consider, for instance, a study by Lehmann and Csank (48), who wished to determine the effects of each of six drugs on various psychological and physiological measures. Responses were studied before and after medication, and the difference in response was attributed to the medication. A control group was not included in the original study, since the investigators believed the responses to be uninfluenced by learning or by other confounding factors.

After the original study, the investigators, apparently feeling that their conclusions were open to question, decided to check their assumption that the change in response following medication was due to the medication alone. They therefore repeated the study, using in place of the previous six drug groups a single control group, subjects for which were presumably selected after completion of the experiment proper.

Although a control group was introduced to avoid confounding drug effects with the effects of such factors as learning, drug effects nevertheless managed to remain confounded, since bias had not been ruled out in assigning subjects to the control group. The structure of this experiment was im-

paired, once the opportunity was lost to assign subjects to drug and control groups at random, from a common pool of applicants. Subsequent efforts to bolster the experiment, no matter how vigorous, could not undo the effects of the investigators' original decision regarding the design of the experiment.

This design shortcoming is by no means atypical. It reflects a common difficulty in distinguishing clearly between design features (such as randomization) which *must* be incorporated into a scientific experiment, and design features (such as the use of baseline measures obtained before medication) which serve only to increase the efficiency of experimentation. Psychopharmacologists have too often discarded *essential* design features in their experiments, while retaining *economic* design features. In view of the serious consequences of failure to discriminate economic from essential considerations, it seems worthwhile to distinguish the two kinds of considerations more precisely.

The following considerations are *essential* to sound experimentation:

1. It is essential to eliminate or to segregate the effects of any factor which introduces bias into the experiment.

2. To avoid confounding drug effects with the effects of factors such as learning, it is essential that observations obtained under the influence of a given drug form be compared with observations obtained under control or standard conditions.

3. If results are to be generalized to a given population, it is essential that subjects be representative of that population. (Representativeness implies the use of a randomizing procedure in the selection of subjects; a strict, rather than haphazard, randomizing procedure must be employed.) Because of practical difficulties in sampling those populations which are of greatest interest to the investigator, the representative sampling requirement is usually more relevant to the interpretation of findings than to the design and conduct of the experiment. (Since the subjects of any two studies are rarely representative of the same population, the results of different studies are, strictly speaking, rarely directly comparable to each other.)

4. It is essential to include a randomizing operation in the assignment of subjects to treatment conditions; a strict, rather than haphazard, randomizing procedure must be employed.

5. In order to obtain an estimate of experimental error, it is essential to include a minimum of two observations for each dosage form. The minimum number of observations *for the experiment as a whole* depends on the kinds of confounding which are to be permitted or excluded from the experiment. Decisions about the kinds of confounding to be permitted or excluded rest in part on economic considerations. A decision to obtain more than the required minimum number of observations depends entirely on economic considerations.

6. The analysis of variance is based on several assumptions which must, theoretically, be satisfied (15). Considerable departures from the assumptions can occur in practice, however, without appreciable error (10). Remedial measures are, moreover, often applicable, where departures from the assumptions are excessive (3). The fulfillment of the assumptions underlying the analysis of variance may therefore be regarded as a quasi-essential feature of studies in which the analysis of variance is employed. Similar considerations apply to special designs, such as the Latin square (which assumes zero interactions; see preceding section) or the usual form of covariance analysis (which assumes linearity of regression).

The following considerations relate to *efficiency* of experimentation:

1. Some designs are more efficient than others. The relative efficiency of two designs depends on the designs themselves (the

Latin square is more efficient than randomized blocks, for example) as well as on the circumstances surrounding a given study.

2. Experimental error can be reduced by physical manipulation of various sources of error. Regulation of error by laboratory controls may be expensive or difficult to achieve in certain instances.

3. The fewer the subjects selected from a given population, the smaller is the likelihood of a given drug effect being detected, and the less precise is the estimate of that effect. The number of subjects selected is usually limited, however, by considerations of time, energy, and money. There is, furthermore, a "law of diminishing returns" relating to subjects: a given increase in the number of subjects is more effective when the original number of subjects is small than when the original number of subjects is large.

4. For the detection and estimation of drug effects, the investigator's energies are more richly rewarded by an increase in the number of subjects selected than by a corresponding increase in the number of observations obtained per subject.

5. Experimental error can be reduced by the use of accessory information, such as baseline measures obtained prior to medication. This accessory information can be utilized in various ways (see above), the relative merits of which are discussed by Cox (12). The collection of accessory information is recommended only when this information is distinctly related to the drug effects being studied, and when the effort and expense of collecting the accessory information are less than the effort and expense of examining additional subjects, or of manipulating various sources of error by physical means. When, on the other hand, subjects are available in adequate numbers, but for limited periods of time, methods of reducing experimental error which do not involve accessory information are indicated.

Before leaving this topic, it is worth noting one more important distinction between essential and economic design considerations. Whereas the investigator can permit himself little leeway in the application of the essential considerations, *each* of which must be satisfied without regard to the others, he may employ one or another economic measure to achieve a given reduction in experimental error. The economic measures operate in concert, competing with one another for inclusion in a given design. Application of the essential considerations depends largely on the investigator's knowledge of these considerations, and on his willingness to apply them rigorously. Application of the economic considerations depends, on the other hand, on the investigator's capacity to appreciate the many psychological, pharmacological, and other facets of the experimental situation, each in proper perspective, and on his capacity to muster his resources to handle, simultaneously and in as judicious a manner as possible, the various practical problems facing him.

Application of the economic considerations is further complicated by another "law of diminishing returns," distinct from the "law" mentioned above, which relates only to number of subjects. The present "law" derives from the fact that economic measures may not only complement one another, but may also provide alternate means to the same end. As economic measures accumulate in a given experiment, there is an increasing likelihood that any newly-introduced measure will duplicate the functions of one or more pre-existing economic measures, and will therefore make no new contribution to the experiment. The investigator should therefore be aware of the precise functions which each economic measure is expected to serve, so that he may best judge the degree to which a newly introduced measure is likely to improve the experiment as a whole.

PSYCHOLOGICAL CONSIDERATIONS AFFECTING EXPERIMENTAL OUTCOMES

SUGGESTION

Suggestion is a process that may readily alter the effects of physical treatments administered to human subjects. No one is more aware of this problem than the physician, who has traditionally assigned sugar pills to patients who need to be reassured by some material token of the physician's continuing concern for their welfare. It is not surprising, then, that the placebo—an inert "dummy" substance bearing a superficial resemblance to the active drug being studied—has become closely identified with attempts to control suggestion, in studies on the psychological effects of drugs or of other physical treatments (43).

The attitude of the subject toward the experimenter may introduce bias into the results and increase experimental error. In most experimental situations, the subjects desire to forward rather than to impede the aims of the investigation (64). Their wish to comply with rather than to resist the supposed effects of the treatment is likely to favor spurious positive results.

Experiments designed to manipulate the effects of drugs on sensory thresholds, by means of suggestion, indicate the kinds of distortion that can be introduced by suggestion. When a hand lotion was applied to subjects by Gault and Goodfellow (21), for the ostensible purpose of decreasing chafing, tactile sensitivity was found to be unaffected. The tactile sensitivity of other subjects (to whom the lotion was introduced as a recently developed sensitizing agent) was found to increase. When alcohol was used to remove the lotion from the latter group of subjects, ostensibly for the purpose of returning tactile sensitivity to normal, the threshold was found to rise.

In an experiment by Goodfellow (23) on coffee drinking, in which subjects were not aware that decaffeinated and regular coffee were being administered on different occasions, auditory thresholds were altered in a similar way by both drinks. After subjects were informed that decaffeinated coffee was being used "just to prove that it was the caffeine that depressed the threshold," all thresholds returned to normal.

Precautions are ordinarily taken, in experiments on drug effects, to reduce the effects of suggestion. Systematic errors due to suggestion can be reduced if the conditions affecting suggestion are kept approximately the same for control and treated subjects. It is customary to administer a placebo to control subjects. While untreated as well as placebo control groups may be employed in a single study, as a further check on the effects of suggestion, this appears to be an uneconomical procedure, since negligible differences have been observed, to date, between untreated and placebo groups (33, 34, 50, 60, 61).

The administration of a placebo to control subjects is one aspect of the "double-blind" procedure. This by now classical feature of drug studies, in which information regarding the drug treatment assigned to a given subject is withheld from both examiner and subject, is designed further to reduce the effects of suggestion. Knowledge of the treatment assigned to a particular subject may influence the examiner as well as the subject. While some examiners might "bend over backward" to avoid inducing the subject to produce the desired results, there would undoubtedly be many instances in which spurious positive results would be favored by the examiner's knowledge of the treatments assigned to particular subjects. "Blindness" on the part of the examiner reduces the temptation (conscious or otherwise) to alter the atmosphere of the examination, the manner of presentation of test procedures, or the quantification of subjective impressions.

The effectiveness of the double-blind procedure may be vitiated somewhat by drug

effects which provide clues, to the subject or to the examiner, regarding the treatment assigned to the subject (1). Placebos employed for research purposes are generally made up to resemble the active substance in appearance, taste, and smell. If the subject has had experience with the active substance, however, and knows or suspects that placebos are being employed in the investigation, he may be led to suspect (in the absence of expected subjective drug effects) that he is the recipient of a placebo. Consideration should therefore be given to the use of placebos which duplicate or mask at least some of the subjective effects of the drug, provided that these effects are incidental to the drug effects being investigated (1). Thus Mead (52), studying the effects of ethyl alcohol on performance, succeeded in masking at least the gustatory and gastric sensations following administration of the alcohol, by adding peppermint oil to both the control and alcohol doses (also see 32, 55). The investigator would also do well to avoid announcing or acknowledging the use of a placebo, and to employ subjects who are unfamiliar with the use of placebos (44).

Observations pertaining to the validity of the double-blind procedure were made in a study comparing the effects of placebo, caffeine (approximately equivalent to that in two cups of coffee), and two strengths of ethyl alcohol (approximately equivalent to that in two martinis and in four martinis) (55). Although the examiner, working "blind," was unable to distinguish between placebo, caffeine, and "two-martini" subjects, on the basis of their gross behavior, he was able to distinguish subjects in these three groups from the "four-martini" subjects. The double-blind procedure thus appears to be of uneven value, being more useful with some dosage forms than with others. It becomes all the more important, then, that the examiner institute measures to increase his "blindness"; that he pay attention to his desires regarding the outcome of the experiment, so that he can prevent these desires from interfering seriously with the experiment; and that he strive for the utmost objectivity in the conduct of the examination. The investigator would also do well, in reporting the results of a study, to indicate the extent to which the double-blind procedure is likely to have achieved its objective in that study.

Additional means are desired for reducing the effects of suggestion on experimental error (the above procedures are primarily designed to reduce bias). Jellinek (40) and Beecher *et al.* (5) have indicated that the inclusion in the experiment of individuals who react to placebo can increase experimental error appreciably, and, conversely, that discrimination among drugs can be sharpened by screening out "placebo reactors." The findings of Lasagna *et al.* (46) suggest that, in comparing the analgesic effects of drugs on postoperative pain, the investigator might well avoid the use of subjects who behave in an immature, dependent, and outwardly responsive fashion. It may also be possible to correct experimental findings for individual differences in suggestibility, by means of the analysis of covariance, provided that appropriate measures of suggestibility can be obtained, and provided that the effects of suggestion are independent of the experimental treatments.[5]

[5] A special form of suggestion may give rise to error when subjects are examined in groups. A subject's need for achievement may be stimulated or depressed when he can compare his performance with those of others (when, for example, he can see that others work more rapidly than he). Communication among subjects should therefore be minimized, unless the investigator is specifically interested in the influence of group processes on drug effects.

Recent findings (58) suggest that certain drug effects may be altered radically by the atmosphere of the group in which the subject finds himself. Group conditions also appear to modify response to medication in animals (8, 29). If it should be established that drug effects can be modified by group conditions, it would be of interest to learn how this modification is brought about.

CONFOUNDING OF THE PSYCHOLOGICAL EFFECTS OF DRUGS

A *given* psychological effect of a drug may be masked by psychological phenomena other than suggestion. This section is devoted to problems of confounding that arise in studies of the effects of a drug on *specified* psychological functions.

The Test Situation. It is important for the investigator to be clear about the kinds of drug effects that he wishes to study. Having a clear idea of the psychological functions which are of primary interest to him, the investigator is able to distinguish the dependent variables of the investigation from those psychological effects which are of secondary interest to him, but which may interfere with the conduct of the investigation. Once the dependent variables are specified, it remains for the investigator to select indices representing these variables. A response is said to be an index of a given psychological function when variations in that response are correlated with variations in the function in question.

Responses which are determined largely by a single function are said to be related *univocally* to that function (28). A univocal relation between response and function is established and maintained by means of appropriate constraints, both external and internal to the subject. The most obvious of these constraints is provided by the experimental task with which the subject is confronted. Appropriate instructions are provided, in addition, to guard against misunderstanding of the task. The constraints provided by the task and by the instructions are not sufficient to guarantee that the subject will actually be engaged by the task. Some subjects approach the experimental task motivated by strong inner needs to perform to the best of their ability. Other subjects are strongly motivated to be helpful to the examiner; when the inner need to excel is insufficient to engage such subjects in the task, they may yet be engaged by means of persuasion, approval, monetary reward, or feedback of results from the examiner.

A test score represents a given psychological function only if minimal "boundary conditions" are met. The response ceases to be a measure of the given function when a minimum of cooperativeness and motivation on the part of the subject is lacking. These are not the only boundary conditions that must be satisfied to preserve the univocal relation between a response and the function it purportedly represents. It is also presupposed that the subject is able to understand the instructions, to perceive the task, and to communicate his responses. Thus, it is assumed that a person can see and write, when a pencil-and-paper test of abstract reasoning is administered to him. But if the subject's ability to receive or transmit information is impaired, his test score cannot be taken to represent his capacity to reason abstractly.

Confounded Effects. The experimenter's attempts to establish the boundary conditions for a given test may be undermined by the very experimental treatment that is being applied. A drug may so alter boundary conditions within the subject that responses to the test may not represent the psychological function in question. Thus, a test may be invalidated by a drug-induced reduction in the subject's cooperativeness or motivation, or by a blocking of the subject's channels of communication.

Several steps can be taken to deal with the problems associated with the confounding of the psychological effects of a drug.

1. Dosage strengths which are likely to produce undesirable confounding effects in a substantial number of subjects should be avoided whenever possible.

2. A drug may block channels of communication which, while utilized in a given test of a function, are not essential for the exercise of the function itself. When the relation between a test and the function it represents is disrupted by such a blocking of communication, the test should be re-

designed, if possible, to utilize channels of communication which remain open. Thus, a test involving visual presentation of stimuli and graphic response might possibly be converted to auditory presentation and vocal response, should a given dose of a drug impair vision and handwriting, but leave speech and hearing unimpaired.

3. When the measurement of drug-induced changes in capacity is affected by drug-induced changes in attitude, the examiner might wish to conduct the examination of each subject in a manner designed to elicit the best performance of that particular subject. Treating each subject in such an individual manner involves subjective judgments on the part of the examiner, who should therefore be wary lest his personal attitudes toward the subject serve to obscure the actual drug effects. Such individualized treatment of subjects is even more risky in the study of drug-induced changes in attitude than it is in the study of drug-induced changes in capacity.

4. The confounding of drug-induced changes in capacity by drug-induced changes in attitude can be reduced by careful selection of subjects. Given a drug which reduces the subject's desire to please the examiner, for example, there is less likelihood of such confounding with subjects having a strong inner need for achievement than with subjects whose strivings are prompted by the desire to gain approval from others.

5. When the confounding of drug effects cannot be avoided, one of two procedures may be followed:

The first procedure is simply to eliminate those subjects in whom the confounding effects occur. This is practical only when the confounding is readily detected, and when it does not occur in too large a fraction of the experimental sample. In studying the acute effects of one of the phenothiazine derivatives on abstract reasoning, for example, one might wish to eliminate those subjects who find it difficult to take a paper and pencil test because their vision has been blurred by the drug.

A second approach centers on attempts to reveal the confounding of drug effects by an analysis of the data. The regression of test performance on the confounding variable may be studied, if a measure of the confounding variable is available. The analysis of covariance cannot, however, be employed to correct for the confounding of two or more effects of the same experimental treatments (26). Factor analysis might be employed to separate the confounded effects; a properly selected test battery must be administered, however, if the variance associated with the secondary, disturbing factors is to be segregated from the variance associated with the direct action of the drug on the dependent variable.

LEARNING

Learning is another psychological factor which can distort the measurement of the psychological effects of drugs. Learning may be a source of difficulty when response sets are progressively altered as the subject becomes more familiar with the task, or when specific stimulus elements of a pretreatment test are recalled during posttreatment testing.

Progressive alteration of response sets may give rise to several problems:

1. Various investigators, *e.g.*, Fleishman and Hempel (18, 19, 20), have shown that the psychological functions underlying test performance may shift as response sets are altered during the course of a test. Such a shift in psychological functions reduces the likelihood that a test score will be related univocally to a given psychological function.

2. There is an increase in experimental error when practice effects during a test differ from subject to subject.

3. Drug effects under study become confounded with some of the practice effects, if response sets are altered during the posttreatment testing as well as during the pre-

treatment testing, and if learning is affected by the experimental treatments.

4. Drug-induced changes in learning capacity may readily be confounded with the drug effects which are of primary interest to the investigator, when crossover or Latin square designs are used. In the crossover design, for example, half of the subjects are first tested in the drug-free state, while the remainder is first tested in the drugged state. The results of such a crossover study may be biased (41); for example, the drug may impair the capacity of subjects who are first tested in the drugged state to understand the task which is set before them.

These problems relating to the response sets of the subject can be minimized when the investigator limits himself to bright subjects or to easy tests. It is desirable, in any event, to make sure that subjects receive adequate test instructions, that subjects work sample problems to test their understanding of the instructions, that subjects learn about their performance on the sample problems, and that sufficient experience be gained prior to the test proper so that subjects are operating near the flat portion of the practice curve during the test proper (41). These suggestions are particularly important when crossover or Latin square designs are used.

Learning may adversely affect the experimental results in yet another way. Post-treatment performance on verbal tests (such as tests of memory for meaningful material) and on projective tests may be distorted because the subject remembers the specific stimulus elements of the task, or the responses to the task, from the pretreatment period. Subjects may remember the earlier material in different degrees, and memory of the earlier material may be affected by the experimental treatments. Alternate test forms which are equivalent or parallel should be provided, if the post-treatment testing takes place within hours, weeks, or possibly even months of the pre-treatment testing (28).

The reports of von Felsinger *et al.* (67) and of Mackworth (50) illustrate how some of the problems relating to learning can be handled.

PHARMACOLOGICAL CONSIDERATIONS AFFECTING EXPERIMENTAL OUTCOMES

The psychological effects of a drug are influenced by the subject's pharmacological history as well as by a host of physiological factors intervening between the experimental administration of the drug and the response of the subject to the experimental task. Some of these influences are discussed below.

SECONDARY DRUG EFFECTS

There are many ways of classifying the effects of a drug. The drug effects which are of primary interest to the investigator are distinguishable from those of secondary interest, once the investigator specifies the dependent variables of the study. Physiological as well as psychological effects may be considered secondary in a given investigation. A most important class of secondary drug effects, in almost any investigation of the psychological effects of drugs in human beings, are those (toxic) effects producing serious physical illness or death (24).

The non-toxic secondary effects may be classified on the basis of their interference or non-interference with the investigation. Aside from the production of serious illness or death, secondary effects may interfere with the investigation in several ways:[6]

1. They may alter the expected course of drug effects by interfering with the absorption or metabolism of the drug. Thus, pylorospasm induced by an oral dose of a drug may hamper or prevent further absorption of the drug (30).

[6] "Behavioral toxicity" is a term referring to the undesirable psychological effects of a drug administered for therapeutic purposes. "Behaviorally toxic" effects do not, however, coincide with any of the classes of drug effects discussed in this paper, which are defined in relation to particular experimental studies.

2. They may prevent the realization of the boundary conditions required for given interpretations of psychological test scores. This has been discussed in an earlier section.

3. They may interfere with investigations of the mechanisms of drug action. One may wish to correlate the psychological and physiological effects of a drug which acts directly on the central nervous system, in order to form inferences relating behavior to central nervous action. Such inferences may be invalid, however, should the drug affect the central nervous system indirectly, as well as directly. Thus, mild distress resulting from direct irritation of the gastric mucosa by the drug may distract the subject from a psychological task, confounding drug effects mediated directly by the central nervous system with effects mediated by the gastric mucosa. This confounding is not a problem in investigations of the *overall* effects of a drug, but it is troublesome when the investigator focusses his attention on a *given kind* of drug action.

SHORT-TERM VERSUS LONG-TERM EFFECTS

The acute effects of a single dose of a drug may differ qualitatively as well as quantitatively from the chronic effects of repeated doses.

Investigations of the psychological effects of drugs on normal human subjects are typically concerned with the acute effects of a single, isolated dose. In these investigations, subjects are generally examined after the drug has been absorbed, but before it has been metabolized. For drugs such as reserpine, which produce their psychological actions after the drug has been metabolized, the investigator will usually wish to study the delayed effects of the drug.

The short-term effects of a single dose may be affected by a history of repeated use of the drug prior to the start of the experiment (24, 41). Goldberg (22) found, for example, that a given drug effect could be produced in abstainers by a lower blood alcohol concentration than that required for chronic drinkers. In studying a drug whose effects are modified by habituation, the investigator may wish to limit himself to subjects having a specified amount of experience with the drug. He can, alternatively, control degree of experience by means of a factorial design. Adjustment for degree of experience, by means of covariance analysis, is a third possibility—but note that placebo groups, or other treatment groups whose performance is unaffected by habituation to the drug in question, cannot be included in the covariance analysis. Since information relating to habituation must usually be based on the subject's testimony, this is one more reason for having reliable subjects, particularly when there is a stigma attached to accurate reporting of the subject's experience with, say, a narcotic drug.

A series of doses may be administered by the investigator, in place of a single dose, as when a graduated increase in dosage is required to bring the internal concentration of, say, a bromide up to a desired level; or, in clinical studies, when little therapeutic effect is expected from a single dose; or in order to minimize undesired effects which may be most prominent after the first few doses (as with some tranquilizers). Experimental error may be introduced from many sources when doses are repeated over a long time, unless the subject is institutionalized, or unless some other form of rigorous control over his daily activities is instituted.

CONTROL OF EFFECTIVE DRUG CONCENTRATIONS

After a drug is administered by one of several routes, it is absorbed, then transported to one or more sites at which it produces its characteristic actions. It is then transported away from these sites, to be metabolized or excreted.

While the experimental treatment consists of the administration of a given dose of the drug, the action of the drug depends on its concentration at the sites at which the given action is effected. There is no one-to-one correspondence between the dosage

level and the effects of the drug at its sites of action, since the intensity of the drug effects depends on the route by which the drug is administered, the pathways in the body along which it travels, the stages in which it is processed in the body, and the rate of processing at each stage (24). Where the experimenter is interested in studying the mechanisms of drug action, he would do well to develop techniques for controlling the concentration of the drug at its sites of action. Such controls are desirable even when there is only a secondary interest in mechanisms of action, since these controls are frequently of value in reducing bias and experimental error. This is especially true in investigations of the acute effects of drugs. Several procedures for controlling the concentration of the drug at its sites of action are considered.

1. Typically, the investigator wishes to study the action of a given dose of a drug. After being absorbed, this dose is distributed among the various body fluids. When a given quantity of drug is administered to different individuals, the drug concentration at a given site of action tends to vary inversely with the overall volume of body fluid. In order to correct for individual differences in the volume of the body fluids, doses are administered in proportion to the weight or surface area of the body (24, 45), either of which is an index of the volume of the body fluids. The administration of drugs on the basis of body weight or surface area results in a substantial reduction in experimental error. (Jellinek and McFarland (41) provide an example of how experimental efforts can be vitiated by a failure to consider body weight or surface area.) When doses are administered without regard to body fluid volume, covariance analysis can be employed, at times, to adjust drug effects for body weight or surface area.

When doses are administered according to body weight or surface area, the experimental treatment can no longer be stated in terms of the absolute quantity of drug administered, since the quantity of drug varies from individual to individual for a given experimental treatment. The experimental treatments are defined in terms referring to the interior of the body: the independent variable is the anticipated concentration of the drug in the body fluids, after absorption is completed. Adjustments for body weight or surface area take into account only the volume of the body fluids; errors due to other "internal" factors, such as variation in the emptying time of the stomach, are unaffected.[7]

2. In studying the acute effects of a drug, the investigator may wish to compare the subject's performance on various tests administered during a single session. Since temporal variations in the drug effects, during the period of test administration, may introduce bias or increase experimental error, it is desirable to maintain the drug effects at a uniform level throughout the testing period. Several methods are available for maintaining drug effects at a uniform level.

Drugs administered orally tend to be absorbed in the small intestine; the temporal course of the drug effects depends, therefore, on the condition of the gastro-intestinal tract. Taken on an empty stomach, a drug tends to be absorbed more readily (30);

[7] In clinical studies dosage is frequently varied according to the examiner's *personal* evaluation of the subject's situation, rather than on the basis of some *objective* criterion such as body weight. Thus, an investigator of the chronic effects of repeated chlorpromazine administration, suspecting that the optimal therapeutic dose differs markedly from situation to situation, may vary the dose from subject to subject, and from time to time for a given subject. This procedure is commendable as a therapeutic technique for use with a drug of established potency. It is, furthermore, clearly appropriate in experiments designed to determine whether individual adjustment of dosage is an effective procedure. Individual adjustment of dosage is of questionable value, however, in experiments designed to evaluate the effects of a drug of unknown potency, for it presupposes that the examiner can detect the very drug effects whose existence is being questioned by the investigator, and whose confirmation awaits the completion of the experiment.

consequently, the rate of rise and decline of the drug effects is more rapid, and the plateau at the peak of the drug concentration curve is of briefer duration.[8] When a drug is ingested after a meal (particularly when the contents of the meal prolong the emptying-time of the stomach), the effects of the drug are somewhat reduced in intensity, but more prolonged and more uniform in time. Quantities of drug which have been metabolized or excreted can be replenished by maintenance doses. Drug levels in the body can be maintained more smoothly, in theory, by the use of sustained release preparations (adequate information regarding the efficacy of these preparations is not yet available, however). For any of those techniques of oral administration, it is important to avoid the use of highly irritating doses, since pylorospasm may prevent the full absorption of the dose, and may produce effects which mask the drug effects under study.

When there is a series of drugs, such as the barbiturates, which produce similar effects but vary in respect to the duration of their actions, the investigator can select a long-acting drug to maintain the drug effects at a uniform level over a prolonged period. The long-acting barbiturates, however, have unpleasant after-effects.

A uniform drug concentration can also be maintained in the blood, over a long interval, by the use of continuous intravenous injection. This method has several distinct advantages (some of which are also enjoyed in the continuous administration of gaseous drugs (66)), as follows:

That component of experimental error resulting from variation in the rate of absorption from the gastro-intestinal tract is eliminated.

An almost instantaneous control of the rate of change of drug concentration in the blood is possible; this is in marked contrast to the situation existing after the oral ingestion of a long-acting drug or of a sustained release preparation.

Undesirable after-effects of a drug are minimized, since the introduction of the drug into the blood stream is readily halted as soon as the experiment is completed.

Sensory cues as to the nature of the drug administered are reduced.

Continuous intravenous administration is not without disadvantage, however. The following observations seem relevant here:

Subjects with inaccessible veins suffer some physical trauma—which can be minimized by care in applying the needle.

Subjects may grow anxious over insertion of the needle, or at the sight of their own blood; care is required to avoid provoking undue anxiety.

Movement of the injected limb is hampered.

Injections lasting much longer than an hour tend to be uncomfortable.

Several investigators (7, 56, 57) have been enthusiastic in recommending continuous intravenous injection, despite its limitations. The method is frequently employed in the clinic, and appears to be worthy of greater attention by investigators planning experimental studies.

Maintaining the concentration of a drug at a uniform level in the blood, or even at the sites of drug action, does not in itself insure that the *effects* of the drug will be maintained at a uniform level (57). Mirsky *et al.* (53), observing that the effects of ethyl alcohol declined even when the blood levels were maintained over a period of hours, postulated a central nervous adaptation (habituation effects occurring within a few hours). The possibility of a short-term adaptation of the central nervous system should be kept in mind when subjects are examined for lengthy periods under the influence of drugs (4).

[8] There are exceptions to this generalization. Drugs such as reserpine have long-delayed effects. The decline of the effects of drugs which are not too readily metabolized, such as methyl alcohol, depends very little on the rate of absorption.

3. There is likely to be some temporal variation in the concentration of the drug at the sites of its action, whatever the mode of drug administration; this variation is, moreover, likely to be correlated from subject to subject. If, furthermore, drug effects vary in time because of a short-term central nervous system adaptation to the drug, this variation is also likely to be correlated from subject to subject. Given such systematic variations in time, a comparison of test scores obtained at widely spaced intervals during a single test session is likely to be biased.

These sources of bias, which are absent in the control group and are therefore unaffected by the use of a control group, may be minimized by randomizing the order in which the psychological tests are administered. Randomization of test order has two disadvantages.

The first such disadvantage is that subjects must be examined singly. The efficiency of experimentation which may be achieved when subjects are examined in groups cannot then be realized.

A second problem is that the internal concentration of the drug, at the time any given test is administered, becomes more variable from subject to subject. If the drug actually affects the functions being studied, there will be an increase in experimental error for the treated subjects; if there is a control group of untreated or placebo subjects, however, there will be no corresponding increase in experimental error for this group. Since procedures for handling heterogeneity of variance may be cumbersome, it is advisable to minimize variations in drug concentration by one of the methods suggested above, even if the order in which tests are administered is randomized. When variations in drug concentration are well-regulated, the increase in experimental error for all treatments combined is likely to be small.

Bias due to temporal variations in drug concentration (or to adaptation) may be greatly reduced, even when subjects are examined in groups, provided that the tests are presented in two opposite orders, each order being employed for one half of the subjects. The reversal of test order removes these sources of bias only when the rate of change of drug concentration (or adaptation) is symmetrical about the midpoint of the testing period. Such symmetry is more likely to exist when sustained release preparations or long-acting drugs are employed, or when drugs are administered after a meal.

The symmetrical placement of equivalent subtests about the midpoint of the testing period also provides a means of minimizing bias due to temporal variations in drug concentration. Such a procedure was followed by Hartocollis and Johnson (32), who compared the effects of alcohol on several tests of verbal fluency. Bias due to temporal variations can be eliminated almost completely by combining the symmetrical placement of equivalent subtests with the presentation of tests in opposite orders to different groups of subjects.

4. The adequacy of the investigator's efforts to control the concentration of the drug in the body fluids can be evaluated, at times, by analyzing the drug content of samples of body fluid. (For many drugs, such as the barbiturates (*cf.* Goodnow *et al.* (25)) and most of the psychopharmacological agents that have excited interest in recent years, satisfactory methods for determining drug levels in small quantities of body fluid have yet to be developed.) The estimate of drug concentration based on a particular sample of fluid may differ appreciably from the concentration at the sites of drug action, especially when the drug is first being absorbed (24, 31). This suggests that the analysis of samples obtained from sites such as the finger-tip which may be remote from the sites of drug action is most meaningful when the drug concentrations of the various body fluids have reached equilibrium. It should be noted that, for a given drug level of finger-tip blood, the effects observed during the early absorption

phase are unlikely to have the same intensity as the effects observed after equilibrium has been reached.

When the internal concentration of the drug varies excessively, from subject to subject, or from test to test, there is a corresponding increase in experimental error. Some of the lost information might be recovered by analyzing the regression of drug effects on drug concentration in the body fluids. (Goldberg's (22) investigations indicate that striking relationships may be observed by studying the regression of performance on blood alcohol level.) Such a regression study may supplement, but cannot be incorporated into, the experiment proper. (Covariance adjustments cannot be made for body fluid level, since the latter is dependent on the experimental treatments (26).)

CONCLUSION

A large number of factors is seen to have a bearing on the design of studies relating to the psychological effects of drugs. In designing a particular study, the investigator takes as his starting point the objectives which he wishes to realize. Which drug effects are to be studied? Which drug comparisons are desired? Is interest focussed on acute drug effects or on chronic drug effects? Is a laboratory study or a clinical study desired? Which kinds of confounding are especially to be avoided; which kinds of confounding can be tolerated?

In order to realize his objectives, the investigator needs an adequate understanding of the many psychological and pharmacological factors which affect the outcomes of studies on the psychological effects of drugs. He should, moreover, appreciate those design features (such as randomization) on which the validity of his conclusions depend, as well as those design features (such as analysis of covariance) which affect the efficiency with which he utilizes the resources available to him. With such understanding, the investigator is ready to formulate procedures capable of eliciting answers to the questions he poses, economically and without bias or undesired confounding.

The investigator's awareness of the considerations which determine the value of a set of drug findings can be heightened further by a study of the strengths and weaknesses of published studies on the psychological effects of drugs. Helpful examples are cited by Jellinek and McFarland (41). The investigator would also benefit from a familiarity with studies such as those conducted at the School of Aviation Medicine (*e.g.*, 34, 36, 59, 60, 61, 62) and by von Felsinger *et al.* (67), which indicate how a number of desirable features can be incorporated into a single design for studying the psychological effects of drugs.

REFERENCES

1. Baker, A. A. and Thorpe, J. G. Placebo response. A.M.A. Arch. Neurol. Psychiat., **78**: 57–60, 1957.
2. Barmack, J. E. The effect of 10 mg. of benzedrine sulfate on the Otis test scores of college students. Amer. J. Psychiat., **97**: 163–166, 1940.
3. Bartlett, M. S. The use of transformations. Biometrics, **3**: 39–52, 1947.
4. Beecher, H. K. Experimental pharmacology and measurement of the subjective response. Science, **116**: 157–162, 1952.
5. Beecher, H. K., Keats, A. S., Mosteller, F. and Lasagna, L. The effectiveness of oral analgesics (morphine, codeine, acetylsalicylic acid) and the problem of placebo "reactors" and "non-reactors". J. pharmacol. and exper. therap., **109**: 393–400, 1953.
6. Bellak, L. *Dementia Praecox.* Grune & Stratton, New York, 1948.
7. Brown, R. R. The order of certain psychophysiological events following intravenous injections of morphine. J. gen. Psychol., **22**: 321–340, 1940.
8. Chance, M. R. A. Aggregation as a factor influencing the toxicity of sympathomimetic amines in mice. J. pharmacol. and exper. therap., **87**: 214–219, 1946.
9. Clark, L. D. Evaluation of the therapeutic effects of drugs in psychiatric patients. Dis. nerv. syst., **17**: 282–286, 1956.
10. Cochran, W. G. Some consequences when the assumptions for the analysis of variance are not satisfied. Biometrics, **3**: 22–38, 1947.
11. Cochran, W. G. and Cox, G. M. *Experimental Designs.* (2nd Ed.) Wiley, New York, 1957.

12. Cox, D. R. The use of a concomitant variable in selecting an experimental design. Biometrika, **44**: 150–158, 1957.
13. Dunnett, C. W. A multiple comparison procedure for comparing several treatments with a control. J. Amer. statis. Ass., **50**: 1096–1121, 1955.
14. Edwards, A. L. *Experimental Design in Psychological Research*. Rinehart, New York, 1950.
15. Eisenhart, C. The assumptions underlying the analysis of variance. Biometrics, **3**: 1–21, 1947.
16. Finney, D. J. *Experimental Design and its Statistical Basis*. University of Chicago Press, Chicago, 1955.
17. Fisher, R. A. *The Design of Experiments*. (4th Ed.) Oliver and Boyd, Edinburgh, 1947.
18. Fleishman, E. A. A factor analysis of intra-task performance on two psychomotor tasks. Psychometrika, **18**: 45–55, 1953.
19. Fleishman, E. A. and Hempel, W. E., Jr. Changes in factor structure of a complex psychomotor test as a function of practice. Psychometrika, **19**: 239–252, 1954.
20. Fleishman, E. A. and Hempel, W. E., Jr. The relation between abilities and improvement with practice in a visual discrimination reaction task. J. exp. Psychol., **49**: 301–312, 1955.
21. Gault, R. H. and Goodfellow, L. D. Sources of error in psycho-physical measurements. J. gen. Psychol., **23**: 197–200, 1940.
22. Goldberg, L. Quantitative studies on alcohol tolerance in man. Acta Physiol. Scand., **5**, suppl. XVI, 1943.
23. Goodfellow, L. D. Significant incidental factors in the measurement of auditory sensitivity. J. gen. Psychol., **35**: 33–41, 1946.
24. Goodman, L. S. and Gilman, A. *The Pharmacological Basis of Therapeutics; a Textbook of Pharmacology, Toxicology, and Therapeutics for Physicians and Medical Students*. (2nd Ed.) Macmillan, New York, 1955.
25. Goodnow, R. E., Beecher, H. K., Brazier, M. A. B., Mosteller, F. and Tagiuri, R. Physiological performance following a hypnotic dose of a barbiturate. J. pharmacol. and exper. therap., **102**: 55–61, 1951.
26. Gourlay, N. Covariance and its applications in psychological research. Brit. J. statist. Psychol., **6**: 25–34, 1953.
27. Gourlay, N. *F*-test bias for experimental designs of the Latin square type. Psychometrika, **20**: 273–287, 1955.
28. Guilford, J. P. *Psychometric Methods*. (2nd Ed.) McGraw-Hill, New York, 1954.
29. Gunn, J. A. and Gurd, M. R. The action of some amines related to adrenaline. Cyclohexylalkylamines. J. Physiol., **97**: 453–470, 1940.
30. Haggard, H. W., Greenberg, L. A. and Lolli, G. The absorption of alcohol with special reference to its influence on the concentration of alcohol appearing in the blood. Quart. J. Stud. Alcohol., **1**: 684–726, 1941.
31. Haggard, H. W. and Jellinek, E. M. *Alcohol Explored*. Doubleday, Garden City, N. Y., 1942.
32. Hartocollis, P. and Johnson, D. M. Differential effects of alcohol on verbal fluency. Quart. J. Stud. Alcohol, **17**: 183–189, 1956.
33. Hauty, G. T. The effects of drugs upon the components of hand steadiness. USAF Sch. Aviat. Med. Proj. Rep., Proj. 21-1601-0004 (Rep. No. 5), 1954.
34. Hauty, G. T. and Payne, R. B. Mitigation of work decrement. J. exp. Psychol., **49**: 60–67, 1955.
35. Hauty, G. T., Payne, R. B. and Bauer, R. O. Effects of normal air and *dextro*-amphetamine upon work decrement induced by oxygen impoverishment and fatigue. J. pharmacol. exper. therap., **119**: 385–389, 1957.
36. Hauty, G. T., Payne, R. B. and Bauer, R. O. Effects of oxygen and dextro-amphetamine upon work decrement. USAF Sch. Aviat. Med. Proj. Rep., Rep. No. 56-127, 1957.
37. Hill, A. B. The clinical trial. J. Med., **247**: 113–119, 1952.
38. Hollister, L. E., Traub, L. and Beckman, W. G. Psychiatric uses of reserpine or chlorpromazine: Results of double-blind studies. In *Psychopharmacology*, N. S. Kline, Ed., pp. 65–74. American Association for the Advancement of Science, Washington, D. C., 1956.
39. Imboden, J. and Lasagna, L. An evaluation of hypnotic drugs in psychiatric patients. Bull. Johns Hopkins hosp., **99**: 91–100, 1956.
40. Jellinek, E. M. Clinical tests on comparative effectiveness of analgesic drugs. Biometrics Bull., **2**: 87–91, 1946.
41. Jellinek, E. M. and McFarland, R. A. Analysis of psychological experiments on the effects of alcohol. Quart. J. Stud. Alcohol., **1**: 272–371, 1940.
42. Keats, A. S., Beecher, H. K. and Mosteller, F. Measurement of pathological pain in distinction to experimental pain. J. Applied Physiol., **3**: 35–44, 1950.
43. Lasagna, L. Placebos. Sci. Amer., **193**: 68–71, 1955.
44. Lasagna, L. The controlled clinical trial: Theory and practice. J. chron. dis., **1**: 353–367, 1955.
45. Lasagna, L. Drug effects as modified by aging. J. chron. dis., **3**: 567–574, 1956.
46. Lasagna, L., Mosteller, F., von Felsinger, J. M. and Beecher, H. K. A study of the placebo response. Amer. J. Med., **16**: 770–779, 1954.
47. Lasagna, L. and von Felsinger, J. M. The volunteer subject in research. Science, **120**: 359–361, 1954.
48. Lehmann, H. E. and Csank, J. Differential screening of phrenotropic agents in man: Psychophysiologic test data. J. clin. exp. Psychopath., **18**: 222–235, 1957.

49. LINDEMANN, E. AND MALAMUD, W. Experimental analysis of the psychopathological effects of intoxicating drugs. Amer. J. Psychiat., **90**: 853–879, 1934.
50. MACKWORTH, N. H. Researches on the measurement of human performance. Med. Res. Council Spec. Rep. Ser., Rep. No. 268, 1950.
51. MCNEMAR, Q. On the use of latin squares in psychology. Psychol. Bull., **48**: 398–401, 1951.
52. MEAD, L. C. The effects of alcohol on performances of different intellectual complexity. J. gen. Psychol., **21**: 3–23, 1939.
53. MIRSKY, I. A., PIKER, P., ROSENBAUM, M. AND LEDERER, H. "Adaptation" of the central nervous system to varying concentrations of alcohol in the blood. Quart. J. Stud. Alcohol, **2**: 35–45, 1941.
54. MOSES, L. E. Non-parametric statistics for psychological research. Psychol. Bull., **49**: 122–143, 1952.
55. NASH, H. Psychological effects of ethyl alcohol and caffeine. (In preparation.)
56. NEWMAN, H. W. Alcohol injected intravenously. Some psychological and psychopathological effects in man. Amer. J. Psychiat., **91**: 1343–1352, 1935.
57. NEWMAN, H. W. *Acute Alcoholic Intoxication.* Stanford University Press, Stanford University, 1941.
58. NOWLIS, V. AND NOWLIS, H. The description and analysis of mood. Ann. N.Y. Acad. Sci., **65**: 345–355, 1956.
59. PAYNE, R. B. Some psychological factors governing the impairment of verbal retention by cerebral depressants. USAF Sch. Aviat. Med. Proj. Rep., Rep. No. 55-52, 1955.
60. PAYNE, R. B. AND HAUTY, G. T. The effects of experimentally induced attitudes upon task proficiency. J. exp. Psychol., **47**: 267–273, 1954.
61. PAYNE, R. B. AND HAUTY, G. T. Effect of psychological feedback upon work decrement. J. exp. Psychol., **50**: 343–351, 1955.
62. PAYNE, R. B., HAUTY, G. T. AND MOORE, E. W. Restoration of tracking proficiency as a function of amount and delay of analeptic medication. J. comp. physiol. Psychol., **50**: 146–149, 1957.
63. PEARSON, R. G. AND BAUER, R. O. The effects of morphine-nalorphine mixtures on psychomotor performance. USAF Sch. Aviat. Med. Proj. Rep., Rep. No. 55-137, 1956.
64. ROSENTHAL, D. AND FRANK, J. D. Psychotherapy and the placebo effect. Psychol. Bull., **53**: 294–302, 1956.
65. SIEGEL, S. *Nonparametric Statistics for the Behavioral Sciences.* McGraw-Hill, New York, 1956.
66. STEINBERG, H. Selective effects of an anaesthetic drug on cognitive behavior. Quart. J. exp. Psychol., **6**: 170–180, 1954.
67. VON FELSINGER, J. M., LASAGNA, L. AND BEECHER, H. K. The persistence of mental impairment following a hypnotic dose of a barbiturate. J. pharmacol. and exper. therap., **109**: 284–291, 1953.
68. VON FELSINGER, J. M., LASAGNA, L. AND BEECHER, H. K. Drug-induced mood changes in man. 2. Personality and reactions to drugs. J.A.M.A., **157**: 1113–1119, 1955.
69. Clinical therapeutic trial of a new drug. Bull. Johns Hopkins hosp., **85**: 221–230, 1949.
70. Conference on therapy: How to evaluate a new drug. Amer. J. Med., **17**: 722–727, 1954.

NEUROLOGIC DISORDERS IN PSYCHIATRIC INSTITUTIONS[1]

FRED A. METTLER, M.D. AND ARCHIE CRANDELL M.D.[2]

THE PSYCHIATRIC HOSPITAL AS A PRINCIPAL SOURCE OF CLINICAL RESEARCH MATERIAL

An accurate understanding of the precise medical condition of the large number of patients in large psychiatric hospitals has always been an obvious prerequisite for any intelligent approach to inquiries into the actual purposes served by these hospitals in the community and for any estimation of what the qualifications of their professional staffs may be. More recently, with the emergence of the public psychiatric hospital as a major source of clinical material for large-scale medical research, it has also become necessary to have accurate medical data on the current status of such patients in order to be able to follow with profit each of the three essential steps in such research —planning, pursuit and evaluation. The value of the research publication written without due regard for the condition of the patient as a whole (and this includes his medical as well as his psychiatric condition) and without regard for his social position and environment is, with the exception of a limited field in which such considerations are irrelevant, relatively slight. When, however, such necessary factors are taken into consideration incipient authors are likely to be appalled by the labor required to come to such an understanding and are likely either to ignore the fact that it is absent or to take the position that it would be irrelevant.

For individual research projects it is not, of course, necessary to evaluate the organic and social situations of all the background material from among which test cases have been selected. It cannot however be denied that even in such restricted situations considerable light may be cast upon the frame of reference in which one is working and one can hardly fail to become aware that the populations in our psychiatric hospitals are selective in many respects other than their psychiatric condition. For it is here that significant sociologic complexes such as have their nucleus in the nature of the law in action, and biologic categories such as congenital malformations, become emergent. One such ancillary selective effect is operative with regard to neurologic disorders and it is with this alone that we shall be concerned in this communication.

DIFFICULTIES WITH EXISTING STATISTICS ON NEUROLOGIC CONDITIONS IN MENTAL INSTITUTIONS

In the course of various psychosurgical studies conducted by ourselves and associates in the past decade, we have frequently had a need for neurologic statistics concerning mental hospital populations, which figures seem to be generally unavailable and difficult if not impossible to compute from the accessible raw data. It is apparent that the intelligent planning, not only of neurologic but also psychiatric research in a mental institution, and the construction of meaningful estimates of the efficacy of therapies must depend upon a knowledge of what type of neurologic material exists there. In physiologic, pharmacologic and pathologic investigations it is absolutely essential to have reliable ideas of how frequent certain neurologic disorders may be. Quite commonly such conditions, if recorded at all in psychiatric hospital records, are treated as secondary diagnoses and are not accessible by electronic sorting

[1] Based in part on an address delivered by Dr. Mettler at the invitation of the N.J. State Bureau for Research in Psychiatry and Neurology, Princeton, New Jersey, April, 1958.

[2] Departments of Anatomy and Neurology, College of Physicians and Surgeons, Columbia, and The New Jersey State Hospital, Greystone Park, New Jersey.

techniques applied to the principal cards. Another serious problem arises from the fact that cases which are admitted to psychiatric hospitals with functional diagnoses often undergo gradual organic changes which are not reflected in the records. A case admitted as a schizophrenic may gradually develop into a fairly obvious heredodegenerative disorder (Huntington's chorea, for example). Since the change is gradual, and the case and its peculiarities are familiar, the diagnosis of record may never be changed. Again, many a schizophrenic patient ultimately develops cerebrovascular disorder but, still being schizophrenic, there seems to be no impelling reason to make any addition to the record and such a case would not be included with the "chronic brain syndrome" categories even though the ultimate reason for hospitalization may depend upon an organic neural defect.

There is another difficulty with the available neurologic statistics on cases in psychiatric hospitals. Social pressures demand that conditions not merely be labelled but that the label be acceptable primarily from a legal and secondarily from a psychiatrically theoretical point of view, neither of which may give a clear medical picture of the case. The crux of the matter consists in the fact that residence in an institution which accepts "mental" cases is not always the equivalent of a complete, current medical diagnosis of psychiatric disorder.

AVAILABLE DATA ON INCIDENCE OF NEUROLOGIC DISORDERS IN MENTAL INSTITUTIONS

Of first admissions to such hospitals in the United States in 1954, approximately the same number of cases bore a diagnosis of chronic brain syndrome (33,801) as of functional psychotic disorders (33,823), according to the *Statistical Abstract of the United States* for 1957. Allowing for the fact that hospital admissions also include about an equal number of additional cases in which are many of personality disorder, and for the fact that the cerebral arteriosclerotic and syphilitics included in the "organic" category exhibit psychopathologic manifestations demanding primary care, it is still apparent that a very considerable proportion of the mental hospital population is in need of competent neurologic attention. Moreover, the functional-psychotic-first-admission is a mobile patient who will probably be out at the end of the year. The same cannot be said for the organic case. There is, as will be seen below and despite the qualification of admissions as "first," a selective tendency, not detectable from "first admission" statistics, which is working toward the concentration of organic disorders in the public "psychiatric" hospitals. While neurologic categories which contribute to the "organic" cohort are not directly determinable from any available statistics, we can still arrive at some opinions concerning them by examining the usage and meaning involved in existing detectable groups which have been developed upon a very broad frame of reference. These groups, if we rely on total hospitalization figures (although percentages for 1933 are used here there is no reason to consider it likely that any considerable shift has occurred meanwhile in percentages, except with regard to syphilis), would seem to be about as follows (13), in terms of the nationwide mental hospital population:

Congenital or developmental disorders	5–6%
Cerebral arteriosclerosis	4–5%
Convulsive disorders	3–4%
Alcohol intoxication	2–3%
Syphilis	less than 2%

Relatively little reliable information is available about the frequency of the remaining potentially significant categories of disorders; conditions such as acute encephalitis, encephalitic residuals, Huntington's chorea, demyelinizing disorders, trauma, tumors, degenerative diseases of the neuro-

muscular apparatus, heredofamilial degenerations, industrial and other poisonings and less frequent neurologic or undiagnosed conditions. Taken together, these must all add up to an additional 15–20 per cent of the nationwide mental hospital population for a reported total of from 30–40 per cent of "chronic brain syndromes" in that population. If to this we add symptomatic manifestations, or secondary conditions such as nerve injuries, in patients with primary functional diagnoses, the total figure for neurologic derangements would seem to suggest that half or more of the chronic psychiatric hospital population is not without neurologic defect.

The widest kind of disparity obtains with regard to estimates of the relative frequency of the smaller individual diagnostic categories. In the case of brain tumor, the national compilation of figures for total mental hospital populations yields an annual incidence of less than one tenth of one per cent. But figures quoted by neurosurgeons from individual hospitals "may range from 3½ to 13½ per cent" (12). Pool and Correll admit that these figures are several times higher than those found in a general hospital. The latter are usually given as between 1–2 per cent (14). It might be supposed that more tumors are present in mental hospitals but are missed. No doubt some, especially those which are metastatic, are but it is questionable if the discrepancy is as large as is implied by Pool and Correll's quoted estimates. It is, for example, rather unusual to find brain tumors in cadavers dissected in Departments of Anatomy which obtain their specimens from mental institutions, even if one looks for such growths for many years. It is only a truism to state that the key to such discrepancies lies in the kind of material attracted by the nature of the service one is talking about and that one must also remain alert to whether the autopsy statistics are based on an across-the-board approach or are highly selective.

In Germany the "psychiatric clinics" contain organic, neurologic, as well as "functional" cases. Of these Zülch says, "Collections of cases from psychiatric clinics are high in frontal lobe tumors, with meningiomas of the olfactory groove, the sagittal sinus, and falx prevailing; collections from chronic mental hospitals, on the other hand, abound in brain tumors which produce early seizures but do not lead to increased intracranial pressure until later (oligodendrogliomas). Malignant glioblastomas predominate in pathological institutes but rarely remain in neurosurgical clinics" (14).

One must not only anticipate a type of tumor incidence in recent admissions to psychiatric hospitals different from that in chronic populations, but also from that encountered in a general hospital. Since cases such as medulloblastomas, spongioblastomas, ependymomas, and perhaps craniopharyngiomas, are extremely rare in psychiatric hospitals for adults, it is theoretically possible for the incidence of tumors of the nervous system in patients in mental hospitals to fall *below* that in general hospitals.

The matter of the nature of a service is important in other respects. Figures below the national percentages can be expected for convulsive disorders in mental hospitals situated in areas where other special hospitals devoted to epileptics exist and, of course, congenital and developmental disorders tend to concentrate in institutions for pediatric defectives when such institutions are available.

Sharp discrepancies may also exist among individual units within mental hospitals with regard to particular disorders. It is well known that the highest concentrations of convulsive disorders are usually to be found in the female admission services (which receive, among other types, those due to renal damage, the sinus thromboses of pregnancy and even relatively recent puerperal toxemias) and in the male chronic

wards (where the traumatized alcoholics with venereal infection are apt to collect). Such differences are likely to result not merely in clashes of opinion as to what the true incidence of convulsive disorder ought to be considered to be in a mental institution, but also lead to considerable confusion as to the precipitating causes of convulsive disorders. Persons familiar with chronic male populations are likely to be of the opinion that alcohol not only greatly increases the probability of convulsions in neurologically traumatized cases but that chronic alcoholism may of itself be an effective cause. (The present writers regard the "rum fit" (9) as a concomitant of edema, except insofar as alcoholism is complicated by trauma, arteriosclerosis or gliosis, secondary to petechiae produced during the convulsions of cerebral edema.) Persons more familiar with female populations are likely to regard alcoholism as less likely to eventuate in convulsions than in polyneuritis. One reason for these descrepancies seems to lie in the differences in the drinking habits between drunken males (who are likely to careen from one bar-and-grill argument to another) and females (who sediment down at home safe from head injury but who neglect the preparation of food as well as their other duties).

THE NATURE OF THE RELATIONSHIP BETWEEN NEUROLOGIC AND PSYCHIATRIC DISORDERS

We have stated above that there is a tendency for neurologic cases to be relatively common in psychiatric hospitals. Although no one will be inclined to take issue with this statement, there is a tendency to discount the facts that, insofar as many of these cases would not be in such institutions at all if they had not become neurologically ill, there is a causal relationship between the two conditions. Moreover, it is clear that the care of some of the acute and many of the terminal neurologic conditions is, *de facto*, a prime social function of these hospitals.

Although it is widely admitted that there is a high degree of concurrence between neurologic and psychiatric conditions it is often necessary to emphasize this by recalling some of the many specific neurologic disorders which are frequently accompanied by psychiatric findings. They are:

Developmental defects.

Generalized infections of the neural system (bacteremias, parenchymatous syphilis, encephalitis, meningitis, complications arising in the course of generalized infections with viruses such as poliomyelitis).

Localized infections or parasitic infestations involving the cerebral cortex, or producing expanding lesions (brain abscess, tuberculomas, cysticercus cysts).

Allergic phenomena (a severe example is Rh factor disease with kernicterus).

Trauma resulting in lacerations of the brain or hemorrhage, concussion or embolism.

Poisoning, including food (ergotism, methyl alcohol), industrial (lead, carbon monoxide, beryllium) and drug (barbiturate, cortisone, chlorpromazine) poisonings.

Nutritional deficiencies.

Cardiovascular disturbances of a nonspecific (Stokes-Adams' syndrome) or specific (cerebral arteriosclerosis, carotid aneurysm) nature (from one point of view of the location of the primary vascular pathology which produces the neural disability).

Metabolic disturbances (premature aging, the lipoidoses, anemias, jaundice, asphyxia, endocrine disorders).

Heredodegenerative disorders of the neural system (Huntington's chorea, Hallervorden-Spatz syndrome, Kalinowsky syndrome).

Convulsive disorders.

Neoplasms and other tumors.

It is often stated, since psychiatric disorder is not invariably associated with the above-mentioned conditions, that when an

association does occur it is of a purely coincidental nature. The neurologic diagnosis may not be searched for or may be ignored, with the result that there may be no way to determine it from the hospital record.

Admittedly, in order to get beyond the possibilities of mere coincidentalism, it is necessary to establish that one or more of the following conditions exist.

1. *It should be possible to demonstrate that there is a higher incidence of psychotic disorder in persons with the particular neurologic disease under investigation than exists in the general population.* This is possible in many of the above-mentioned conditions.

2. *The neurologic disorder in question should exist more frequently in psychotics than in normal persons.* While there is less likelihood of normal or psychotic persons developing neurologic complaints than there is of neurologic cases developing psychotic phenomena no adequate data exist for determining whether psychotic persons are more vulnerable to the subsequent appearance of particular neurologic phenomena than are normals.

3. *The nature of the relationship between the neurologic and psychiatric disorder ought to be of a fairly characteristic type.* This is true of conditions such as meningitis and delirium, encephalitis and behavior disturbances, and several other constellations.

4. *The occurrence of the psychiatric and neurologic phenomena should both depend upon the involvement of particular structures or functions, or both should be manifestations of a common underlying pathologic process.* Such a demonstration is conditioned by one's definition of what an essential or characteristic psychotic phenomenon may be. Most physicians, would, for example, include hallucinations but not aphasic disturbances in that class. How shall we consider the "inaccessible," "strange," "uncontrolled," "rigid," or pathologically forgetful individual in whom no hallucinative processes can be demonstrated? When such phenomena appear in the course of a neurologic disorder are they to be looked upon as basic disturbances in the psychic process or as "reactive" expressions? Since there seems to be a growing tendency on the part of neurologists to regard "behavioral" conditions as a part of their field of legitimate interest, a rather high correlationship is being developed between neurologic and psychiatric manifestations in such areas as may be encompassed by a broad approach to the convulsive disorders. Whatever position we adopt on this question it is apparent that, if the association between certain neurologic and psychiatric conditions is of more than a coincidental nature, the one may have precipitated the other or they may both be due to a common cause, in which case what we know about the structural substrate of neurologic conditions may shed light on psychopathology. The outstanding possible exception to this rule derives from the fact that it is still far from clear whether some of the psychiatric disturbances encountered in neurologic complaints are not (like paranoid behavior in failing audition) primarily reaction patterns revealing a vulnerable underlying personality quite apart from the particular neurologic disability. This was Bleuler's argument in connection with schizophrenic manifestations, such as paranoia, occurring in the course of cerebral arteriosclerosis. Still, inquiry into such questions ought to shed some light upon the problem of what portions or functions of the neuraxis cannot be infringed upon without precipitating psychiatric symptoms, providing one can come to agreement as to what symptoms are incontrovertible proof of a psychotic process.

The problem is a difficult one and one which we shall pursue in another connection elsewhere. For the present it is worth making explicit what has been implied above, namely that since we are aware that psychiatric and neurologic disorders are frequently associated we need to know

which are the most common patterns of association. Then there is need to search in both areas to see if such an association does exist, and not merely content ourselves with either a neurologic or psychiatric categorization.

SURVEY OF THE GREYSTONE POPULATION

In the fall of 1955, and before "tranquilizing" drugs had been introduced into general use in the New Jersey State Hospital at Greystone Park (cases which had been so treated were placed at that time in one ward and no cases with pre-existing extrapyramidal signs were admitted to it), a survey was conducted there with a view to determining the prevalence of the major neurologic categories, especially paralysis agitans, in such a representative mental institution. (A study (4) published in 1951 of the 25 mental hospitals in the United States containing 4000 or more patients in 1949 had disclosed that the N. J. State Hospital at Greystone Park was a thoroughly representative large public civil psychiatric hospital.)

On June 30, 1955, Greystone contained 6853 patients (considerably more than the 5698 figure for 1949). Table 1 presents their distribution in the categories noted on page 149, according to hospital records.

The following other diagnostic categories could be identified: Epidemic encephalitis, 32 cases (0.4 per cent); brain trauma, 32 cases (0.4 per cent); senile brain disease, 56 cases (0.7 per cent); tumor, 2 cases (less than 0.1 per cent, or the same as the nationally reported figure) and mental deficiency, 394 (5.9 per cent). There was no way to determine from the records alone the number of cases in some individual conditions, such as multiple sclerosis, but these were expected to have been included in a category of 37 cases designated as "diseases of unknown and uncertain cause." There were two additional cases listed as "chronic brain syndrome of unknown cause." It was not possible to identify separately such symptomatic states as chorea or choreoathetosis, nor to determine what proportion of the cases had symptomatic states such as hemiplegias, paralysis agitans or peripheral nerve injuries.

TABLE 1

Greystone Patient Distribution, Selected Categories, June, 1955

Diagnosis	N	% Total Population
Congenital or developmental disorders	9	0.1
Cerebral arteriosclerosis	966	14.6
Convulsive disorders	104	1.5
Alcohol intoxication	197	2.9
Syphilis of the nervous system	271	4.1

Because of the time lag involved in making an actual survey in order to see what these records meant, and because of mobility in and out of as well as within the hospital, a cut-off date of November 5, 1956, was accepted as the day of prevalence, new neurologic cases being picked up to that date and all listed cases not then in the hospital being dropped from the roster.

Before examining the manner in which the cases found compared with the available listing, it is necessary to emphasize that no individual psychiatric hospital can be compared directly with the figures for the national prevalence of neuropsychiatric disorders, for hospital populations are not true cross-sections of the national distribution. There are, moreover, individual peculiarities in every institution. In the case of The New Jersey State Hospital at Greystone Park, there are few if any pediatric cases and no criminal psychotics except statutory sex offenders. An unusual feature of the hospital, since the thirties, has been an influx of patients from the New York Neurological Institute. Certainly the prevalence of previously treated neurologic cases with very full records is high. This has been due to the fact that residents of New Jersey, who are treated in the New York Neurological Institute (which being situated on the East

Bank of the Hudson draws a not inconsiderable proportion of its cases from New Jersey), are not eligible for state hospital care in New York. More recently the referral system between the counties and state hospitals in New Jersey has been altered and a smaller proportion of Neurological Institute cases may be expected to appear at the New Jersey State Hospital at Greystone Park in the future, but at the time of the survey it was possible to follow the protracted course of a considerable number of types of neurologic cases which ultimately appear in psychiatric institutions admitting the chronically ill. Epileptics form the largest such group. Former neurosurgical cases constitute a small but steady and spectacular influx.

One important feature of the survey which immediately became apparent was the fact that the orientation of the resident staff at Greystone toward neurology was in the direction of the dramatic. There was a distinct tendency to consider, and not entirely without justice, cerebrovascular disease and its chronic manifestations as conditions in which neurologists are not interested. Even epileptics were frequently passed over as presumably of no interest. The effect of this orientation (for which neurologists have themselves to blame because of a relative disinterest in commonplace, chronic conditions), is to favor the impression of the staff that mental hospitals contain fewer cases of neurologic disorders than even the statistical registers show must be in them. Thus, the June 1955 statistics from Greystone indicated that there were at least 1744 cases which were classifiable as organic (excluding the alcoholics and special groups other than mental deficiency) but rapid sampling by the hospital staff of a segment of 1426 of the more chronic patients resulted in the selection of only slightly over 300 cases which the physicians in charge considered to display neurologic disorders. After examining all of these a neurologic diagnosis could be verified (though not necessarily the original one) in 299 cases. A striking feature of this group was that it contained a very low percentage of cases exhibiting cerebrovascular disease other than those cases with rather startling findings (such as pontine hemorrhages), definite full-blown hemiplegias or meningovascular syphilis. In a population of one thousand one would expect between 50 and 60 cerebral arteriosclerotics but the group of 299 contained fewer than ten patients who fulfilled criteria noted below. The anticipated 50–60 so-called "senile" cases were absent. They were present, of course, in the parent sample of 1426 cases, but it is not to be expected that a psychiatric staff will immediately think of such cases as neurologic and it may be doubted if the private practitioners of neurology would look upon them with any insatiable degree of professional enthusiasm.

With a psychiatric hospital population of about six thousand patients one would expect between 1800 and 2400 patients to present evidence of organic disease of the nervous system. In the present instance all the subsequent definitive examinations were conducted by the senior author, since it was felt that to rely upon examinations employing variable techniques and criteria would introduce avoidable confusion into the work. Yet many difficulties and obscurities remained. It takes a long time to see 6000 patients, especially when they are being moved about in a large hospital, and some leave and others enter. There is, moreover, the factor of change in the patient's condition itself. Limitations as to the degree of thoroughness with which examinations are to be conducted have to be accepted and it is frequently found that histories and records are incomplete, inconclusive or have to be accepted at face value even though suspect. As matters stood we came to the cut-off day with a roster of 1192 "brain syndrome or deficiency" cases with evidence

TABLE 2

Greystone Organic Disease Data Compared with Anticipated Figures and 1955 Records

"Organic" Category	Anticipated	1955 Records	Found
Congenital or developmental disorders	285–341	9	196
Cerebral arteriosclerosis	228–285	966	124
Convulsive disorders	171–228	104	292
Alcohol intoxication	118–171	197	20
Syphilis	less than 118	271	132
Other	855–1140	122	428

of neurologic disorder out of a total of 5704[3] against an expected figure between 1711 and 2281. Table 2 presents a comparison of this survey data with the figures anticipated in terms of the nationwide population mentioned above, and with the hospital statistics for June, 1955, at the start of the survey.

In the tabulation of "found" cases, Table 2, there is some duplication in that the same patient may be listed in more than one category. The first revealing discrepancy to be observed is the low prevalence of so-called alcohol intoxication. It is obvious that such a diagnosis is essentially an acute one and that even cases such as those admitted to the "formal" category—notably neuritis and Korsakoff's syndrome—may actually be deficiency disorders. The only explanation for the high figures on alcohol intoxication in a resident population (rather than on admission) is that the diagnosis given at admission has not later been changed. The durability of the admission diagnosis introduces a very serious hidden error in all mental hospital statistics.

The second principal discrepancy, that in cerebral arteriosclerosis, stems from our personal difficulty in coming to a decision as to what criteria are really acceptable for the diagnosis. We are by no means certain that even the 124 cases listed above as "found" would show pathologic evidence of diminished cerebral blood flow with structural arterial change in the middle cerebral artery. No doubt many of the cases rejected as arteriosclerotic have some structural alteration in the vessels but did it in fact produce a diminution in blood flow? Unquestionably some cases would show cerebral atrophy, but cerebral atrophy can appear without cerebral arteriosclerosis. In the present instance the only cases admitted to the "found" category were those exhibiting some evidence of at least a temporary, localizing lesion, in the presence of presumptive evidence of cerebral arteriosclerosis. This procedure coincides closely with the implications of the recently published "Classification and Outline of Cerebrovascular Diseases" (7). There is a great need for extensive, careful correlation between clinical and pathological findings in cases of supposed cerebral arteriosclerosis.

In the category of other "found" neurologic disorders the principal diagnoses were:

Aphasia	20
Apraxia	9
Athetosis	8
Brain Injury	79
Brain tumors, unoperated	19
postoperative	5
Cerebral degeneration	16
Chorea, postencephalitic, tubercular, luetic or arteriosclerotic	13
Huntington's	8
Choreoathetosis	6
Encephalitic residuals (various types) (14 with parkinsonism)	37
Hemiparesis, hemiplegia	65
Hemorrhage cerebral, nontraumatic, other than with enduring hemiparesis or hemiplegia	25
Multiple sclerosis	7
Nerve peripheral, injury of	22
Paralysis agitans and parkinsonism	47
Pituitary insufficiency	4
Pluriglandular syndrome	5
Pseudobulbar palsy	5

[3] Between June 30, 1955 and November 5, 1956 the population in Greystone was reduced as the result of the opening of an additional state hospital in the southern part of New Jersey.

The difference between 400 and 428 was more than made up (because of duplication) by four or less cases each in the following categories: acrodynia, amyotrophic lateral sclerosis, anemia (sickle cell), beryllium poisoning, bulbar palsy, carbon monoxide poisoning, carbon tetrachloride poisoning, combined system disease, heredoataxia (cerebellar), heredofamilial tremor, lead poisoning, meningitis (status post-meningitis), metastatic carcinoma to vertebral column or brain, muscular dystrophy, myoclonus, narcolepsy, optic atrophy, poliomyelitis, polyneuritis (acute infectious) and toxemia of pregnancy.

Mental deficiency, diplegia, Klippel-Feil deformity and the like were included under congenital or developmental defects. Korsakoff's syndrome and avitaminotic (alcoholic) neuritis were included under alcohol intoxication. Not included in any of the above groupings were cases of toxic reactions (including allergies) to drugs (such as dilantin). Also excluded were fainting due to pulmonary or cardiac disease, nonpituitary endocrine disorders (including parathyroid tetany but not pluriglandular syndrome; no Cushing syndrome was found) or reaction patterns—such as overventilation—precipitating neurologic signs. Unless non-traumatic convulsive phenomena were of a repetitive type without specific obvious causes, such as toxemia of pregnancy or cerebral edema (due to "rum fits") or barbiturate withdrawal, they were not included.

The preexisting neurologic diagnosis which it was found most necessary to change was multiple scleorsis and the diagnosis which had been most frequently missed was brain injury.

The most common difficulty in arriving at a diagnosis, aside from the problem of arteriosclerotic brain disease, was brain tumor and the reason for this was not because of its frequency but because of the necessity of utilizing special auxiliary diagnostic skills and apparatus not always to be had.

Convulsive disorders. Although this study was not concerned with treatment it was obvious that the problem of pharmacologic control of convulsive disorders is, both individually and numerically, the most important neurologic therapeutic problem in hospitals of this type. The convulsive group further contains some of the patients most difficult to examine and dangerous to treat. The convulsive category in a mental institution is made up of every conceivable type. It is not uncommon to encounter patients who have self-induced seizures in addition to their involuntary attacks. The self-induced seizure may be a genuine attack produced by deliberately utilizing epileptogenic stimuli or entirely feigned. Some patients present all three varieties at different times (making the regulation of their medication extraordinarily difficult) and, furthermore, their involuntary attacks may be made up of both grand and petit-mal seizures.

Brain trauma. The psychiatric hospital is the location of termination of a number of categories which the clinical neurologist treats for some period during their development. In addition to the primary categories of congenital or developmental disorders (see above) and many cases of paralysis agitans, should be mentioned cases of brain trauma—especially those in whom a diagnosis of alcohol poisoning has been made—and various terminal aspects of partial, unsuccessful, palliative or basically inept neurosurgery.

Cerebral arteriosclerosis. Cerebral arteriosclerosis had been diagnosed by the hospital staff in 14.6 per cent of the total June 30, 1955 Greystone population. This is much higher than the 4–5 per cent reported in the nationwide mental hospital population of 1933 and probably reflects to some degree the trend to be anticipated with a lowering of the death rate in older age brackets and better mobility of younger patients under psychopharmacologic regimens. The same phenomenon has been observed by others

(5) and a doubling of the figures on cerebral arteriosclerosis has occurred in practically all psychiatric hospitals in the 15 year period between 1935–1950. (The rise in this category since 1919 has been about 500 per cent (6)). Adams and Vander Eecken (1) report that in a general hospital population 24 per cent of cases in which the brain was examined exhibited cerebrovascular disease. It is apparent that the rate would be expected to be at least as high in a mental hospital population. Of this group of 24 per cent, approximately one third had atheromatous changes which were manifested by infarcts alone somewhat more frequently than by definite thromboses. The most frequent location of pathologic changes due to atherosclerosis was in the basal ganglia which showed small cavitations. Most such cases had been hypertensive. (Fifteen per cent of the cases of cerebrovascular disease in Adams and Vander Eecken's series showed massive hemorrhages.)

The criteria used in most mental institutions for establishing a diagnosis of cerebral arteriosclerosis are usually similar to those laid down by Murphy (8) who states that it "may be made in an elderly person with signs of arteriosclerosis elsewhere, particularly in the vessels of the scalp and eye, in whom neurologic disability has progressed slowly and inexorably but with episodes of temporarily more severe deficit. Mental and physical symptoms and signs indicate involvement of the brain generally. There are numerous and widespread neurologic abnormalities of greater and less degree with special references to the crainal nerves and pyramidal and extrapyramidal projection systems."

Since Murphy admits that there is no close correlation between signs of systemic or ocular arteriosclerosis and cerebral arteriosclerosis (see also (2) p. 298; "evidence of general arteriosclerosis is generally lacking"), the critical question is whether the clinical signs he describes are necessary concomitants of the pathologic changes of cerebral arteriosclerosis. He does not discuss the possibility that such findings may be caused by one of the primary brain atrophies in which there is no evidence of cerebral arteriosclerosis (1). If we ignore the so-called systematic atrophies, and disallow Peters' (11) conception of idiopathic epilepsy and the endogenous psychoses as degenerative processes, we are still left, following Murphy's criteria, with the problem of how to distinguish between cerebral arteriosclerosis and senile and presenile diffuse atrophies. The premature appearance of memory failure and a personality change in the direction of "*Witzelsucht*," when of a familial nature and rapid downhill course, is usually taken as suggestive of Alzheimer's disease and most clinicians would agree with Brodal (3) that a distinction between senile dementia and presenile sclerosis (Alzheimer's disease) is only possible upon a clinical basis (see also (10), p. 233). But the clinical distinction between senile and presenile primary atrophy does not greatly aid us in distinguishing between atrophy secondary to arteriosclerosis, if this really is common, and primary generalized brain atrophy itself. Ordinarily what psychiatrists seem to do is to decide in favor of the former diagnosis if some of the basic qualities of the individual's personality or habit patterns are preserved and detectable, and in favor of the latter if the degree of disorganization is extreme and characterized by exposure of the underlying personality urges of the individual. The impression is abroad that arteriosclerotic brain disease decomposes the internal ability of the person to monitor his activities without early disorganization of the social amenities, whereas the dementias due to primary atrophy result in an initial erosion of surface traits, such as cultural inhibitions and learned elements such as memory, with exposure of the fundamental character traits of the individual. In the clinical decision the appearance of remissions is taken to provide strong evidence in favor of a diagnosis of

arteriosclerotic brain disease. If one is to profit by Adams and Vander Eecken's experience one would expect evidence of brain damage due to atherosclerotic brain disease in about eight per cent of general autopsies with two out of every five such cases exhibiting evidence of thrombosis as the cause of the infarction. Embolism without evidence of atherosclerosis ought to be just about as frequent. Most of the atheromatous patients should show hypertension and most would have basal ganglia damage. So much for a general population, but we cannot say that the same set of conditions would obtain in a committed population.

The available national statistics on cerebral arteriosclerosis in committed populations are influenced by a spurious element which is always encountered when the scientific evaluation of medical fact is extraneously influenced by social considerations. In this case the social consideration is the law, for if a diagnosis of senile dementia is made, it is a foregone conclusion in most jurisdictions that the courts will deny testamentary or disposing capacity to such an individual and it may become impossible for a wife to manage the affairs of an incapacitated husband without entering upon a long and expensive legal process. On the other hand, a diagnosis of arteriosclerotic brain disease does not preclude the possibility of finding a fortunate "lucid interval." Although it is getting to be practically impossible to protect the testamentary capacity of even entirely healthy and well-oriented individuals of advanced age (a factor contributing to their disappearing confidence in the social order), court practice surrounding the *per vitam* sale of property, owned in common, tends to encourage the discovery of "lucid intervals" and has a devastating effect upon hospital statistics relating to the cerebral dementias and cerebral arteriosclerosis.

It seems worth remarking that we, like O'Leary, Adams, Vander Eecken and many others, were unable to satisfy ourselves that cerebral arteriosclerosis is an inevitable consequence of advancing age. We also obtained the impression that when it does occur it is likely to begin relatively early, suggesting that a metabolic defect may very well be involved. One would suppose that a mental hospital population in which the median age is well over the national figure would be precisely the material to confirm the notion that cerebral arteriosclerosis is a concomitant of advancing years. While it is true that failing mentation appears ultimately to be inevitable, it is by no means clear that those cases in which cerebral arteriosclerosis was most certainly present first developed the condition late in life. In most instances the disorder seems to have already been apparent before the age of 60. (Social custom seems to be retiring people at the wrong age; if they get to 65 without difficulty it would appear that they are likely to be able to keep going much longer). This situation is thrown into relief when one fixes one's attention on a particular manifestation such as paralysis agitans (and we suspect it would also be true if other indicators were selected), but the critical data would seem to be in those homes for the aged which select only healthy older persons and which subsequently witness little if any cerebral arteriosclerosis. This area is importunate for inquiry and so is one related to it which Professor Kolb has emphasized, the development of methods for the early detection of cerebral arteriosclerosis and for measuring its progress.

CONCLUSIONS

1. The available statistics on psychiatric hospital populations which appear in the customary available sources, yield little information of value to the neurologist. Social considerations, as well as medical and scientific fact, influence the choice of diagnostic designations in a not inconsiderable number of instances. Since preference is of necessity given to a psychiatric diagnosis, secondary neurologic diagnoses may be entirely in-

accessible. Data as to actual hospital content are difficult to obtain in comparison with statistics on first admission, and the latter are weighted in favor of mobile categories. First admissions figures give no reliable indication of the nature of the cases in the durable hospital population. When diagnostic statistics are available as to diagnoses on total resident population, they may reflect past rather than present conditions and some categories, such as "cerebral arteriosclerosis," may be clinical conjectures rather than true etiologic diagnoses. At best they leave one guessing about the functional condition of a patient who may be quite able to care for himself (as in a geriatric institution) or may be a bed-ridden, paralytic dement.

2. The largest group of "organic" cases in large psychiatric institutions for all classes of patients is the *complex* composed of cerebrovascular disorders and primary cerebral degenerations.

3. The most frequent neurologic *problem* demanding perspicacious neurologic attendance in such institutions is posed by the convulsive disorders.

4. The neurologic *disorder* most frequently missed is brain trauma.

5. The encountered neurologic *diagnosis* most frequently found to be in error (other than cerebral arteriosclerosis and alcoholism) is multiple sclerosis.

6. Among the additional neurologic problems are congenital and infectious disorders of the nervous system, industrial and other poisonings (including drug poisonings), heredodegenerative disorders, nutritional disturbances and peripheral nerve injuries. Brain tumor is a minor but spectacular problem in mental institutions and one which is of greater significance in new admissions than in the chronic population.

7. There is a close relation between neurologic disease and psychiatric disorders. There are at least two reasons for this. Psychiatric reactions are frequently precipitated by the stresses imposed by neurologic disorders, and psychiatric and neurologic disorders are frequently produced by the same causes. Either manifestation may precede the other or they may occur concurrently.

REFERENCES

1. Adams, R. D. and Vander Eecken, H. M. Vascular diseases of the brain. Ann. Rev. Med., **4**: 213–252, 1953.
2. Brain, W. R. *Diseases of the Nervous System.* 3rd ed. Oxford Univ. Press, London, 1945.
3. Brodal, A. On the anatomy and pathogenesis of cortical brain atrophies. Acta psychiat. et neurol. scandinav., Suppl. **74**: 140–163, 1951.
4. Kramer, M. Critique of a standard psychiatric population. In *Evaluation of Psychosurgery. Proc. Third Res. Conf. Psychosurgery, 1951,* Overholser, W., ed., pp. 120–130. (Public Health Service Publication no. 221) U. S. Dept. Health, Education & Welfare, Washington, D. C., 1954.
5. Malzberg, B. A statistical study of patients in the N. Y. civil state hospitals, April 1, 1950. Psychiat. Quart., **26**: 70–85, 1952.
6. Malzberg, B. A comparison of first admissions to the N. Y. civil state hospitals during 1919–1921 and 1949–1951. Psychiat. Quart., **28**: 312–319, 1954.
7. Millikan, C. H. et al. A classification and outline of cerebrovascular diseases. Neurology, **8**: 395–434, 1958.
8. Murphy, J. P. *Cerebrovascular Disease.* Year Book Pub., Inc., Chicago, 1954.
9. New York Academy of Medicine. Committee on Public Health Relations. *Alcoholism as a Medical Problem,* Kruse, H. D., ed., p. 83. Hoeber-Harper, New York, 1956.
10. O'Leary, J. L. Ageing in the nervous system. In *Problems of Ageing,* Cowdry, E. V., ed., pp. 223–238. 3rd ed. Williams and Wilkins, Baltimore, 1952.
11. Peters, G. *Spezielle Pathologie. Der Krankheiten des zentralen und peripheren Nervensysteme.* Thieme Verlag, Stuttgart, 1951.
12. Pool, J. L. and Correll, J. W. Psychiatric symptoms masking brain tumor. J. M. Soc. New Jersey, **55**: 4–9, 1958.
13. Zubin, J. and Scholz, G. C. Regional differences in the hospitalization and care of patients with mental diseases. Pub. Health Rep., Suppl. **159**: 8–9, 1940.
14. Zülch, K. J. *Brain Tumors, Their Biology and Pathology,* p. 58. Springer Pub. Co., Inc., New York, 1957.

The following two papers are being published in order to acquaint Western clinicians and investigators with some of the work and thinking currently in progress in the U.S.S.R. It is hoped that these papers will have special value and interest despite the difficulty of checking certain references and of establishing ready communication with the authors.

The Editors

ARTERIOSCLEROTIC PSYCHOSES[1]

A Review of the Foreign Literature From 1940 to 1956

V. M. BANSHCHIKOV AND G. V. STOLIAROV

Progress in modern medicine, as stated by a number of foreign authors, has led in many countries to a considerable increase in the average duration of life and, as a result of this, to an increase in the percentage of persons of advanced age. Thus for example, according to the data of Dublin (25), the average duration of life in the U.S.A. increased from 50 years in 1900 to 64 in 1942, while the percentage of persons over 45 increased between 1900 and 1940 from 17.8 to 26.5 and should reach 40.3 by 1980. A similar shift in the composition by age of the population is taking place in England, Sweden, and other countries. It is only natural that under these conditions the problems of cardiovascular diseases, which for the most part attack persons of advanced age, should become more timely both for medicine as a whole and for psychiatry as well.

According to Dublin, cardiovascular diseases, together with kidney diseases, are the cause of more than half of all deaths in persons 50 years and older. Arteriosclerosis is the most frequent disease of the circulatory system. According to Attinger (9), the number of patients with cardiovascular diseases (of which the most frequent was arteriosclerosis) was greater than the number of patients with tuberculosis, malignant tumors, diseases of metabolism and the digestive system, and psychoses and neuroses combined. Schinz and Reich (64) state that the mortality per 100,000 from arteriosclerosis, excluding coronary arteriosclerosis, rose in Switzerland per 100,000 population from 6.26 in 1901–1904 to 14.79 in 1950–1952. According to Katz and Dauber (41), of 120,000 cases of death from intracranial circulatory disease, cerebral arteriosclerosis was the cause of death in 75,000. R. Brain (15) reports that circulatory disease of the central nervous system was the cause of death of 69,388 persons in England and Wales in 1951 as compared with 56,048 in 1942; this amounts to 14 per cent of all deaths. P. White (71) considers arteriosclerosis to be the most frequent cause of death and disability "except for man's inhumanity to man."

Not unexpectedly, the increase in the number of arteriosclerosis patients led to an increase in the number of patients with arteriosclerotic psychoses. The work of Dayton (23), published in 1940, was based on a detailed statistical treatment of the data on 89,100 patients entering the Massachusetts Psychiatric Hospital from 1917 to 1933. Arteriosclerotic psychoses, according to his data, stood in second place with 11.36 per cent and were exceeded in frequency only by schizophrenia. Per 100,000 popula-

[1] Banshchikov, V. M. and Stoliarov, G. V. Zhur. Nevropatol. i Psikhiatrii **57** (8): 1044–50, 1957. Translation of this paper has been furnished by the National Institutes of Health, Public Health Service, as a part of their Russian Scientific Translation Program.

tion of the same age with a morbidity of not less than 90 per cent of arteriosclerotic psychoses, this was the leading cause of disease among men (1143) and was in second place in women (805), exceeded only by the senile psychoses. The distribution by years showed undeviating growth of arteriosclerotic psychoses both in absolute numbers and per 100,000 population. The corresponding figures in 1917 were 261 and 40, and in 1933 they were 720 and 78. According to Rothschild and Sharp (62), of all the patients admitted to the Worcester, Mass., psychiatric hospital, patients with arteriosclerotic and senile psychoses amounted to 27.2 per cent in 1943 and to 33.7 per cent in 1945. Maltzberg (43) calculated the expected frequency of arteriosclerotic psychoses in New York State per 10,000 individuals born and obtained the following data.

	1920	*1930*	*1940*
Men.....	51.2	113.9	191.6
Women.....	37.4	95.0	185.3

Clow (20) cites the work of Landis and Paige, in whose opinion arteriosclerotic psychoses are the only disease the frequency of which increased conspicuously (536 per cent) in New York State from 1912 to 1936. According to Wright and MacDevitt (78), cerebral arteriosclerosis was in second place among the causes of admission to psychiatric hospitals in the U.S.A. in 1954. However, the mortality and duration of life of patients with arteriosclerotic psychoses has not changed materially in the last 20 years, according to Pollak (52).

It is interesting that the increase in morbidity from arteriosclerotic and senile psychoses is exceeding the percentage increase of those entering the group of the aged in the population. According to a report by Colb (21), the number of persons 65 or more in the USA increased by 57.5 per cent between 1936 and 1951, while the number of first admissions to psychiatric hospitals there increased by 95.3 per cent; about 80 per cent of those admitted were patients with arteriosclerotic and senile psychoses. D. Blasko (cited by Colb) reports that in Connecticut the population increased by 9.6 per cent from 1950 to 1954, the number of persons 60 and older by 14 per cent and the number of persons of this age admitted to psychiatric hospitals increased by 39 per cent.

The general problems of the etiology and pathogenesis of arteriosclerosis which are discussed with such interest in the general medical literature involve psychiatrists almost not at all. Regardless of the fact that the term "arteriosclerosis" itself (introduced, as is known, by Lobstein in 1833) is not generally accepted, the term "arteriosclerotic psychoses" is still used in all psychiatric work.

Since cerebral arteriosclerosis develops, as a rule, in older people (although certain authors (13, 40) note increasing frequency of arteriosclerosis and circulatory disease of the brain in younger age groups), a question arises concerning the interrelation of senile and arteriosclerotic psychoses. The belief that we are dealing here with two different diseases is not subject to doubt. This has been confirmed again by Arab (8) and by Raskin and Ehrenberg (54) on the basis of autopsy and clinical anatomical comparisons. However, Raskin and Ehrenberg, in studying the autopsy and clinical data of 270 patients who died in psychiatric hospitals, come to the conclusion that the combination of arteriosclerotic and senile changes is encountered considerably more often than "pure" changes. Rothschild (60) also refers to the frequency of combined arteriosclerotic and senile anatomical changes but considers that the mixed clinical picture is rare. Critchley (22) notes the frequency of the combination of senile psychoses with cerebral arteriosclerosis. He suggests that the question of the relation of these two diseases is of greater academic than practical interest since during life cerebral arterio-

sclerosis is usually associated with old age. This point of view is in agreement with the fact that in the publications of clinicians arteriosclerotic and senile psychoses are usually considered together.

The question of the interrelation of hypertensive disease and arteriosclerosis is discussed in a number of publications, but the conclusions of the different investigators are often contradictory. Thus while Ker'nei (42) assumes that high blood pressure and arteriosclerosis should be distinguished from each other "although in patients with high blood pressure there is always sclerosis of the intracranial vessels," Eros (26) considers that arteriosclerosis and sclerosis of the large vessels cannot be differentiated. Hypertension, in his opinion, is the most important factor in the development of arteriosclerosis. Botton (14) found that cerebral arteriosclerosis was accompanied by high blood pressure in two-thirds of the cases studied. Hetenyi (35) considers that a large portion of persons suffering from arteriosclerosis are patients with hypertension in the second or sclerotic stage of the disease. In most clinical papers this distinction is ignored and authors either limit themselves to stating that the blood pressure was increased in the patient with a diagnosis of arteriosclerotic psychosis whom they are describing (1, 7, 20, 59), or they consider the increased blood pressure as one of the important symptoms of arteriosclerosis (48). G. Williams (74) supports the necessity of differentiating hypertensive and arteriosclerotic psychoses. In his opinion arteriosclerosis with continuous hypertension displays clinically discontinuous symptoms which fluctuate in intensity, such as headache, dizziness and anxiety; rarely, delirium or anemia is seen. Later the disease becomes chronic, with loss of memory and a decline in initiative and working capacity. A second group consists of patients with sclerosis of the large arteries in whom, in addition to the symptoms previously described, there is early onset of easy fatigability, memory defect, and depressive and paranoid traits. Another result of the disease is degeneration of the personality. Unfortunately, the author does not report the factual data he used for his conclusions. Goodman (32) describes six cases of protracted hypertension with psychotic episodes which produced a picture of paranoia, hallucinosis and even of "complete degeneration," ending, however, in recovery. The degeneration in these cases occurred late. The changes noted at autopsy consisted chiefly in arteriosclerosis and slight thrombosis. This disease should be differentiated from arteriosclerosis. Rosenberg (57) noted that in 17 cases of fatal malignant hypertension, the severity of the arteriosclerosis did not always correspond to the severity of the hypertension. Hughes, Dodgson and McLennan (38) describe 61 patients with "chronic cerebral hypertensive disease" in whom prolonged hypertension was accompanied by emotional and intellectual disturbances and "social indifference," often in combination with pseudobulbar symptoms. However, in these cases the degree of severity of the symptoms and their fluctuations were not correlated with the severity of the hypertension and variations in blood pressure; at autopsy it was discovered that the arteriosclerotic changes were chiefly in the subcortical area. Russek and Zohman (63) noted that the clinical picture of the so-called hypertonic encephalopathy may be observed in the absence of hypertension. In patients with hypertension and arteriosclerosis, lowering the blood pressure to the normal level not only did not improve their condition but often made it worse. The authors conclude that arteriosclerosis is the sole cause of repeated cerebral episodes even when hypertension is present. K. Stern (68) comes to directly opposite conclusions. Analyzing seven cases of disease with a subacute course and episodic psychoses of the organic type, he expresses the hypothesis that "these cases, in which the psychosis is doubtlessly related to the

arteriosclerosis, are cases of genuine hypertension." But "pure" arteriosclerosis is expressed by gradual degeneration of the personality; it is only an occasional trait of pathological aging of the central nervous system and should therefore be associated with the anatomical and clinical picture of senile dementia. Mayer-Gross, Slater and Roth (44) consider that making a separate category of hypertensive psychoses is justified only on the basis of the relation to hypertonic encephalopathy, which is manifested as paroxysmal headaches accompanied by nausea, vomiting, and disturbances of consciousness (including stupor and coma). Personality changes of a neurotic nature are also characteristic of hypertension. Thus the question of the clinical differentiation of arteriosclerotic and hypertonic psychoses cannot be considered as solved.

Arteriosclerotic psychoses are known to have been classified first as an independent group on the basis of autopsy data. In recent years a number of separate papers have appeared, the authors of which attempt to explain the peculiarities of the clinical picture of cerebral arteriosclerosis on the basis of the changes found at autopsy. Thus Eros (26) divides arteriosclerosis into two anatomical types, hyperplastic and hypoplastic. In the former, focal symptoms predominate; in the latter, mental changes more often develop earlier and more severely. In later stages of the disease these distinctions are obscured. Rosenberg (57) divides the clinical symptoms into three groups: 1) symptoms of increased intracranial pressure; 2) symptoms of numerous slight attacks; 3) symptoms of massive foci. However, most investigators emphasize the absence of correlation between the extent and localization of the changes seen at autopsy and the mental condition of the patient. Thus Irish (39) noted the appearance of psychotic symptoms in only 170 of 1000 cases of embolism, thrombosis and brain hemorrhage. In this study the presence and nature of the psychotic changes did not depend on the area of whichever artery was attacked. Rothschild (61), noticing the discrepancy between the anatomical and clinical data, voiced the hypothesis that arteriosclerotic psychoses may be considered as "the manifestation of decompensation on the cerebral level." The possibility of either somatic or psychic compensation depends to a certain degree on the individual personality. A severe sclerosis may cause psychic changes in anyone, but in "poorly balanced" personalities the disease emerges earlier and more lightly. The special characteristics of the premorbid personality which facilitate the development of a psychosis may be quite various. Raskin and Ehrenberg (54) also noted the lack of correlation between the severity of the sclerosis and the degree of degeneration. The lack of correlation between the clinical and autopsy findings was noted by Allen (7), Schlinker (65), Neumann and Cohn (50), and a number of other authors. Zeman (79) considers that the psychic changes are usually caused by the complication of the arteriosclerosis by one or more supplementary factors (fall in blood pressure, anemia, pulmonary disease, etc.). Hofman (cited by Botton (14)) goes even further and comes to the conclusion that disease of the blood vessels of the brain in general plays a small role in the appearance of the so-called arteriosclerotic psychoses.

Only a few papers relate to the study of the circulation and metabolism of the brain in vascular psychoses. Freyhan, Woodford and Kety (30), investigating the velocity of the circulation in the brain and the oxygen requirement of the brain in six patients with arteriosclerotic dementia, noted decreases in both in comparison with normal controls. Brobeil, Harter, Herrman and Kramer (16) established that an increase in the oxygen requirement and an increase in the vascular walls takes place in parallel with the development of cerebral arteriosclerosis. Fazekas, Bessman, Cotso-

nas and Alman (27) investigated changes in the same factors using different functional loads in 30 patients with cerebral arteriosclerosis and obtained data showing that the vessels of these patients preserved the ability to dilate and were sufficiently elastic that they could maintain changes in tonus in answer to one or another of the stimuli, although this reaction was quantitatively less pronounced than in healthy persons. In this paper no description is given of the clinical symptomatology. Fazekas, Kleh and Finnerty (28) investigated 27 patients with senile and arteriosclerotic psychoses. It appeared that the velocity of the circulation in the brain and the oxygen requirement was noticeably decreased in comparison with the normal. The presence of hypertension increased the resistance of the vessel walls but did not noticeably affect the oxygen requirement of the brain or the velocity of the circulation. The authors do not consider it to be proved that the psychic degeneration is a result of these changes. D. Wilson (76) also suggests that the significance of anoxia as a cause of the death of cells and of the psychosis in cerebral arteriosclerosis is only hypothetical, since too many of the factors of cerebral metabolism remain unknown. Carrone (17), determining the urinary 17-ketosteroid content, obtained data in nine cases which were on the lower border of the normal range for the corresponding age group.

The question of the effect of various factors on the appearance of cerebral arteriosclerosis and on its further course is dealt with by a smaller number of papers than in the preceding ten years. Allen (1) Berke (10) and Clow (20) attribute importance to heredity. Of 100 patients with arteriosclerotic psychoses observed by Clow, 33 had relatives with cardiovascular diseases. Other factors noted were constitutional inferiority (26), infection, starvation and other factors leading to increased parasympathetic tonus (66). Allison (3) considers that hypoglycemia and cerebral anoxia usually lead to a reversible state of confusion but may also lead to dementia. Botton (14), Willens (72) and Clow (20) deny that alcoholism plays a role. Clow noted that there was alcoholism in the anamnesis of nine of 100 patients. Willens, comparing the brains of 519 alcoholics and of 600 non-alcoholics who died when older than 35, found that atherosclerosis (*sic*), including the coronary and cerebral types, was encountered considerably more rarely in alcoholics. According to Botton's data, of 220 mentally ill patients in whom no arteriosclerosis was detected at autopsy, 34 were alcoholics, and alcoholics constituted only 28 of a group of 308 patients with severe cerebral arteriosclerosis. Considerable importance in the development of psychoses is attributed by the majority of the investigators to psychic trauma, *e.g.* loss of close relatives, material insecurity, retirement from business (62), emotional tension (7, 10, 20). Stern (68) attributes decisive significance to social factors while Hetenyi (35), on the contrary, denies their role. Wilson (76) considers that absence from business pursuits, deprivation of interests and being placed in a psychiatric hospital, as well as poverty, or the feeling of being personally not needed; all these may facilitate degeneration because they all increase the limitations on the horizon, which in turn lead to a decrease in the circulation of the brain.

The description of the clinical picture and course of arteriosclerotic psychoses in the recent literature add nothing really new to the descriptions given in older papers and psychiatric textbooks. Allen (1) considers a gradual beginning and a slowly progressive personality degeneration as characteristic. However, along with these often insignificant changes, more than half these cases develop an acute psychosis with confusion and sometimes with hallucinations. He emphasizes the frequent occurrence of depression, paranoid symptoms, and display of instincts which were inhibited before the onset of disease, and classifies three basic forms: 1) cerebral

arteriosclerosis without psychosis and characterized by easy fatigue, loss of interests, somnolence and irritability; 2) the "little strokes"; and 3) arteriosclerotic dementia. By "little strokes" Allen, agreeing with other authors, is referring to a special form of cerebral arteriosclerosis pointed out by Alvarez and characterized by the development of a series of repeated thromboses of the small arteries. These thromboses do not cause paralysis, and focal symptoms are usually absent. Often the sequelae of such a thrombosis are manifested only in the patient's mental condition: his memory worsens, and there are changes in character which appear unexpectedly. Repetition of the thromboses leads to degeneration. In the intervals between them the patient's condition remains unchanged, and there may even be some improvement. The disease may last for 10 to 20 years, and the prognosis is unfavorable. Rothschild (59), like Allen, notes the possibility both of slow degeneration which is undeviatingly progressive regardless of fluctuations in its course, and also of fulminating progressive psychoses with confusion which in 50 per cent of the cases end in death. The syndromes of depression and paranoia are often observed, but permanent, systematized delirium is not seen. In 100 cases of arteriosclerotic psychosis, Clow (20) observed that 17 had an acute beginning, while 83 developed gradually. Delirium was observed in 34 patients, confusion in 18, depression in 34 (including 19 with agitated depression), manic states in five, and paranoia in four. Degeneration was noted in 14 patients. Somewhat different data are cited by Beckenstein and Gold (12), who noted an acute onset in 77 of 100 patients seen by them. According to the data of Mayer-Gross, Slater and Roth (44), prolonged depression similar to the endogenous type, is encountered only rarely in cerebral arteriosclerosis, perhaps in 5 to 10 per cent of these cases. Usually the depressions are brief and are accompanied by clouding of consciousness. They observed epileptiform attacks in 15 to 20 per cent of the cases, more often as generalized attacks. M. Roth (58) distinguishes a separate group of all psychoses arising in old age; he places the maniacal and depressive syndromes seen then in a special group of affective disturbances of old age. In Alpers' paper (5) only two basic forms are given: the neurasthenic syndrome and the degenerative picture. Sjögren (67) distinguishes arteriosclerotic dementia and post-apoplectic insanity. Among the clinical criteria making possible a diagnosis of arteriosclerotic psychosis, Williams, Quesnel, Fisch and Goodman (75) include: 1) the presence of focal symptoms; 2) isolated areas of relatively well-preserved memory and association; 3) lability of mood; 4) hypertension; 5) peripheral or retinal sclerosis. Alpert, Bigelow and Bryan (6) also list five diagnostic criteria: 1) subjective sensation; 2) peculiarities of the psychopathological picture; 3) hypertension; 4) systemic and 5) cerebral arteriosclerosis. However, Alpers, Forster and Herbut (4), on the basis of anatomical studies of the retina and cerebral arteries in 100 individuals, came to the conclusion that the absence of retinal sclerosis could not serve as proof of the absence of sclerotic changes in the blood vessels of the brain. In the papers of Roth (58) and of Hopkins and Roth (36), diagnostically important symptoms of arteriosclerotic psychoses are said to be: a) dementia in combination with focal symptomatology; and b) a course with remissions or fluctuations at any stage of the disease if they are associated with such symptoms as irresolution, epileptiform attacks, or preservation of the critical sense.

Some authors have attempted to reveal the individual mental characteristics of atherosclerosis patients by various psychological tests. Delay, Pichot, Dursap and Perse (24) based their investigations on the concept of Babcock, according to which verbal tests appeared to be least hampered in physiological and pathological degeneration. Comparing the data obtained from ex-

aminations of patients with arteriosclerotic and senile dementia, these authors found that "intellect and activity" were preserved to the same degree in both states, but the arteriosclerotic patients performed less well on the verbal tests. They consider the only possible explanation to be the existence in the patient of "latent aphasia" which, in their opinion, may have a definite diagnostic significance. Rashkis (53) compared the characteristics of the fulfilling of various tasks by patients with schizophrenia, bulbar paralysis and cerebral arteriosclerosis. He found that arteriosclerotics compared poorly in tests and were usually not in condition to explain their actions but tried to find justification for this. In these cases if the explanation was acceptable, it was logical. Pinkerton and Kelly (51) noted in psychological examinations an increasing loss of ability to deal with abstract ideas which recalled the picture of the development of this ability in children, but in the reverse direction.

A number of clinical papers contain descriptions of individual cases with unusual manifestations of cerebral arteriosclerosis. Similar cases have been described before. Thus Chavany and Tapta (18) report a case of apraxia restricted only to certain kinds of movements of the fingers. Meyer (45) observed two patients in whom a state of confusion passed off without leaving any changes. Mezei (46) reports a case of cataplexy which changed into narcolepsy, and a case of myoclonia; D. Williams (73) reported a patient in whom arteriosclerosis was associated with uremia; Weber and Blum (70) had a patient with arteriosclerosis the clinical course of which resembled that of epidemic encephalitis.

There are almost no references to the prognosis of arteriosclerotic psychoses in the available literature. Clow (20) reports that of 100 of his patients, 11 were cured, 12 experienced considerable improvement, 31 had some improvement, there was no change in the condition of 30, and six died; in 29 patients in whom the psychosis developed immediately after cerebral insult, only nine improved. The condition of the patients was not followed after discharge. M. Roth (58) collected data on the condition of patients with arteriosclerotic psychoses at six months and again at two years after their admission to the hospital. Six months later eight of 36 patients were discharged, 16 remained in the hospital and 12 died. Two years after admission, of 22 patients three were discharged, three were still in the hospital and 16 had died. Allen (2) states that the prognosis in psychoses caused by sclerosis of the blood vessels of the brain is considerably more favorable than in senile psychoses.

Treatment recommended for patients with arteriosclerosis consists basically of nursing care, diet, small doses of iodine and administration of vitamins; when irritability is present, the usual sedative drugs are given (1, 7, 47, 62). However, there is no full agreement here. Thus Katz (19) denies the usefulness of diet; G. Williams (74) casts doubt on the therapeutic value of iodine. Wilson and Hohman (77) carried out a course of parenteral administration of iodine to two patients and observed no effect at all. G. Williams (74) and Russek and Zohman (63) state that it is useless to lower the blood pressure. In early stages of the disease psychotherapy is recommended (62), "logical reassurance" (47). Allen (1, 2) considers as permissible the administration of electroconvulsive therapy in combination with curare for the treatment of depressions in arteriosclerotic patients. He refers to the predisposition of elderly patients with cerebral arteriosclerosis to attacks of acute excitement, for which he advises treatment without the use of drugs. If drugs are used, their excretion should be hastened by the prescription of laxatives and diuretics, enemas, etc. In the treatment of such patients great importance is attached to the observance of a strictly conservative regime. An odd point of view is voiced by D. Wilson (76). He considers that psychosis or neurosis is a defense against "psychogenic forces" and casts doubt on the usefulness of treat-

ment, especially in depressions, since the patient, as a result of treatment, is "deprived of his last defense." Reese and Kant (55) used aminophyllin successfully for the treatment of arteriosclerotic psychoses. They consider it to be especially indicated in confusion and also in treating headache, dizziness and insomnia. There was improvement in 23 of 32 patients. Garnett and Klingman (31) obtained improvement in seven of nine paients with the beginning symptoms of cerebral arteriosclerosis and senile psychosis by the use of intravenous injections of 1 to 5 ml of a preparation containing 1 mg/ml of cytochrome C for two weeks. The catamnesis (from one to nine months) confirmed the permanence of the therapeutic effect. Forster and Bradford (29) tried to use the same method in the treatment of nine patients with arteriosclerotic and senile dementia, but they did not note any favorable results. Bamford (11) describes 15 patients in whom improvement was achieved by the use of lipotropin or oxytropin (preparations which include chiefly vitamins of the B group and U.S.P.-type thyroid). Herrligkoffer (34) achieved success in the treatment of seven patients with paditine tablets or injections. Huber (37) recommends the use of preparations of nicotinic acid (nikovaza) to decrease or abolish such symptoms as malaise, anxiety and dizziness. Sometimes it was even possible to influence the characterological changes caused by the disease. The attempts by Taylor (69) to treat arteriosclerosis with histamine were without effect: of nine patients not one had an improvement in mental condition, but coronary embolism was noted in four of them. Haibach (33) treated 79 patients with the initial symptoms of cerebral arteriosclerosis with venostasin (chestnut extract containing aesculetin) injections of 4 ml with 0.125 mg of strophanthin (about 20 injections per month).

O. Negri (49) treated patients with senile and arteriosclerotic psychoses with reserpine and with reserpine and Ritalin. Reserpine was given to 40 patients at 0.75 to 2.5 mg daily. It gave the best results in manic states and with querulousness. With depression the results were considerably worse: of ten patients there was noticeable improvement only in two and insignificant improvement in three. The treatment had no effect on the patients' disturbances of memory and cognition. Combined treatment with reserpine (0.5 to 0.25 mg) and Ritalin (5 to 20 mg) three times daily was given to 49 patients. By this method the sedative effect of reserpine was retained, while at the same time the patients became more active and their orientation for time and place improved. The treatment of patients with depression also appeared to be considerably more effective: slight to considerable improvement occurred in eight of ten patients. The side effects of reserpine were also avoided. S. Rizzente (56) used reserpine for the treatment of 45 patients with arteriosclerosis and hypertension; the neurasthenic syndrome was predominant in their clinical picture. Improvement took place in 40 patients; the author relates this to the decrease in blood pressure.

Summarizing the data contained in the reference literature, it may be said that the psychiatric problem of arteriosclerotic psychoses have not been given enough attention. The very important questions of pathogenesis, differential diagnosis, prognosis and treatment are just as far from solution as before. Sheinker's statement (19) of 1946 remains true even today: "Cerebral arteriosclerosis is the most widespread and one of the least understood pathological conditions of the brain."

REFERENCES

1. Allen, E. New England J. Med., **245**: 677–684, 1951.
2. Allen, E. Bull. New York Acad. Med., **31**: 366–375, 1955.
3. Allison, R. Brit. M. J., **2**: 1286–1289, 1952.
4. Alpers, B., Forster, F. and Herbut, P. Arch. Neurol. & Psychiat., **60**: 440–456, 1948.
5. Alpers, J. Ohio M. J., **39**: 1108–1109, 1943.
6. Alpert, H., Bigelow, N. and Bryan, L. Psychiat., Quart., **21**: 305–313, 1947.
7. Alvarez, W. Am. J. Nursing, **47**: 169–170, 1947.

8. Arab, A. Rev. Neurol., **91**: 22–36, 1954.
9. Attinger, E. Praxis, **27**: 476–479, 1947.
10. Baere, K. Zentralbl. ges. Neurol. u. Psychiat., **103**: 375, 1943.
11. Bamford, T. New York J. Med., **51**: 2913–2916, 1951.
12. Beckenstein, N. and Gold, L. Psychiat. Quart., **19**: 398–411, 1949.
13. Berlin, L., Tumarkin, B. and Martin, H. New England J. Med., **252**: 162–166, 1955.
14. Botton, J. Encéphale, **44**: 350–396, 1955.
15. Brain, R. Ann. Int. Med., **41**: 675–681, 1954.
16. Brobeil, A., Harter, O. Herrman, E. and Kramer, K. Klin. Wchnschr., **32**: 43–44, 1954.
17. Carrone, G. Schweiz. Arch. Neurol. u. Psychiat., **74**: 166–178, 1954.
18. Chavany, J. and Taptas, J. Presse méd., **55**: 10, 1947.
19. Circulation, **5**: 98–100, 1952.
20. Clow, H. Am. J. Psychiat., **97**: 16–26, 1940.
21. Colb, L. Am. J. Psychiat., **112**: 627–636, 1956.
22. Critchely, M. In *Problems of Ageing*, Cowdry, E. V., ed., pp. 518–534. Williams & Wilkins Co., Baltimore, 1942.
23. Dayton, N. *New Facts on Mental Disorders*. C. C Thomas, Springfield, Ill., 1940.
24. Delay, J., Pichot, P., Dursap, R. and Perse, J. Rev. Neurol., **83**: 180–191, 1950.
25. Dublin, L. In *Problems of Ageing*, Cowdry, E. V., ed., pp. 91–110. Williams & Wilkins Co., Baltimore, 1942.
26. Eros, G. J. Neuropath. & Exper. Neurol., **10**: 257–294, 1951.
27. Fazekas, J., Bessman, A., Cotsonas, N. and Alman, R. J. Gerontol., **8**: 137–145, 1953.
28. Fazekas, J., Kleh, J. and Finnerty, A. Am. J. Med., **18**: 477–485, 1955.
29. Forster, W. and Bradford, E. J. Ment. Sc., **97**: 718–724, 1951.
30. Freyhan, F., Woodford, R. and Kety, S. J. Nerv. & Ment. Dis., **113**: 449–456, 1951.
31. Garnett, R. and Klingmann, W. Amer. J. Psychiat., **106**: 697–700, 1950.
32. Goodman, L. Arch. Neurol. & Psychiat., **62**: 445–454, 1949.
33. Haibach, A. Deutsche med. Wchnschr., **80**: 1573–1574, 1955.
34. Herrligkoffer, R. Med. Monatsschr., **6**: 353–354, 1952.
35. Hetenyi, G. Acta med. Acad. sc. hung., **4**: 425–438, 1953.
36. Hopkins, B. and Roth, M. J. Ment. Sc., **99**: 451–463, 1953.
37. Huber, H. Wien. med. Wchnschr., **104**: 773–774, 1954.
38. Hughes, W., Dodgson, M. and McLennan, D. Lancet, **267**: 770–774, 1954.
39. Irish, C. Am. J. Psychiat., **96**: 897–903, 1940.
40. J. A. M. A., **134**: 608, 1947.
41. Katz, L. and Dauber, D. J. Mt. Sinai Hosp. New York, **12**: 382–410, 1945.
42. Kern'ei, I. Acta med. Acad. sc. hung., **4**: 455–457, 1953.
43. Malzberg, B. Psychiat. Quart., **19**: 122–138, 1945.
44. Mayer-Gross, W., Slater, E. and Roth, M. *Clinical Psychiatry*. Cassell, London, 1954.
45. Meyer, H. Arch. Psychiat., **184**: 272–273, 1950.
46. Mezei, B. Psychiat. u. Neurol. Wchnschr., **43**: 15–17, 1941.
47. Munk, F. Deutsche med. Wchnschr., **47**: 915–920, 1941.
48. Mutrux, M. Monatsschr. Psychiat. u. Neurol., **125**: 19–38, 1953.
49. Negri, F. Arch. psicol. neurol. e psichiat., **17**: 441–451, 1956.
50. Neumann, M. and Cohn, R. A. M. A. Arch. Neurol. & Psychiat., **69**: 615–636, 1953.
51. Pinkerton, P. and Kelly, J. J. Ment. Sc., **98**: 244–255, 1952.
52. Pollak, O. J. Nerv. & Ment. Dis., **102**: 27–29, 1945.
53. Rashkis, H. J. Nerv. & Ment. Dis., **106**: 650–670, 1947.
54. Raskin, N. and Ehrenberg, R. Am J. Psychiat., **113**: 133–137, 1956.
55. Reese, H. and Kant, F. Am. J. Psychiat., **103**: 713–732, 1947.
56. Rizzente, S. Arch. psicol. neurol. e psichiat., **17**: 454–456, 1956.
57. Rosenberg, E. Arch. Int. Med., **65**: 545–586, 1940.
58. Roth, M. J. Ment. Sc., **101**: 281–301, 1955.
59. Rothschild, D. Am. J. Psychiat., **98**: 324–333, 1941.
60. Rothschild, D. Arch. Neurol. & Psychiat., **48**: 417–436, 1942.
61. Rothschild, D. Am. J. Psychiat., **100**: 501, 1944.
62. Rothschild, D. and Sharp, M. Dis. Nerv. System, **8**: 123–128, 1947.
63. Russek, H. and Zohman, B. New York J. Med., **49**: 1411–1414, 1949.
64. Schinz, H. and Reich, T. Deutsche med. Wchnschr., **80**: 952–954, 1955.
65. Schlinker, J. Am. J. Path., **22**: 565–571, 1946.
66. Schmitz, B. Med. Klin., **12**: 376–378, 1952.
67. Sjögren, H. Nordisk med., **52**: 1083–1091, 1954.
68. Stern, K. Canad. M. A. J., **45**: 513–517, 1941.
69. Taylor, J. Dis. Nerv. System, **8**: 154–155, 1947.
70. Weber, F. and Blum, K. Lancet, **242**: 503, 1942.
71. White, P. J. Gerontol., **6**: 151–153, 1951.
72. Willens, S. J. A. M. A., **135**: 1136–1139, 1947.
73. Williams, D. Brit. J. Radiol., **16**: 142–143, 1943.
74. Williams, G. Cleveland Clin. Quart., **14**: 168–175, 1947.
75. Williams, H., Quesnel, E., Fisch, V. and Goodman, L. Am. J. Psychiat., **98**: 712–715, 1942.
76. Wilson, D. Am. J. Psychiat., **111**: 902–906, 1955.
77. Wilson, W. and Hohman, J. J. Nerv. & Ment. Dis., **118**: 351–354, 1953.
78. Wright, I. and McDevitt, M. Lancet, **26**: 825–830, 1954.
79. Zeman, F. J. Mt. Sinai Hosp. New York, **17**: 1075–1082, 1951.

CYBERNETICS AND CERTAIN MODERN PROBLEMS OF PHYSIOLOGY OF THE NERVOUS SYSTEM[1]

B. KH. GUREVICH

Cybernetics is literally the science of control (from the Greek, meaning the art of control). Interest in it among the scientific world has not diminished since the moment of publication of Norbert Wiener's *Cybernetics or Control and Communication in Animal and Machine* in which the general principles of cybernetics were first formulated. This interest, it is understood, does not betoken an unquestioning acceptance of all the theses of the author, but so greatly without opposition has it been, nonetheless, that certain propositions advanced by cybernetics merit a many-sided discussion.

The ideas of cybernetics in many ways recall theories of physiology which have already been worked out, or models observed in neurophysiology and applied in the area of the regulation of functions in the organism. Well known are the contributions to the science of control and connections in the living organism of such outstanding physiologists as Claude Bernard, Sechenov, Cannon, Starling, Sherrington, Magnus, and also a large number of contemporary neurophysiologists. By the admission of Wiener, the conceptions of Pavlovian "nervism" and the teachings of I. P. Pavlov on the conditioned reflexes had, and continue to have, exceptional importance in principle for the growth of the ideas of cybernetics. Hence the interest which physiologists and representatives of contiguous branches of knowledge are displaying in cybernetics is understandable. In this article an attempt is made to review briefly certain problems of nerve physiology in relation to the basic positions of cybernetics.

[1] GUREVICH, B. KH. Vestnik Akad. nauk SSSR (Herald of the Academy of Science, USSR), **27**: 31-40, 1957. Translation of this paper has been furnished by the National Institutes of Health, Public Health Service, as a part of their Russian Scientific Translation Program.

In living organisms, the investigator is dealing with intact, self-regulating systems. According to cybernetics, the essence of the principle of self-control (auto-regulation) lies in the fact that the movement and action of large masses or the transmission and transformation of large amounts of energy, are directed and controlled with the aid of small masses and inconsiderable amounts of energy, bearing information in the form of knowledge and commands received by the system in connection with the fulfillment of the action.

In this definition it is easy to distinguish certain separate elements, or divisions, of cybernetics: the theory of auto-regulation from the point of view of energy relations and thermodynamics, the theory of information, and finally, the theory of the inverse connection. It is the inverse connection, that is, the receiving of knowledge and commands in connection with the performance of action, which actually makes the system auto-regulative and allows, with respect to the principles of control, some comparison between artificial mechanisms and the living organism.

In technology, one of the precursors of such mechanisms with inverse connections is the centrifugal regulator or governor of the steam engine constructed by Watt. About the axis, linked with the shaft of the machine, revolves a device with spheres which, upon an increase in the amount of steam admitted and in the swiftness of revolution, separate in such a way that there is a diminution in the steam admitted and in speed of revolution. Hence the steam engine regulates itself, or is self-governing with respect to speed.

Watt did not realize that, having created his regulator, he had accidentally discovered a universal principle of auto-regulation.

At present the principle of negative inverse connection is applied in numerous automatic devices.

In the area of the regulation of the internal environment, the concept of compensation and of inverse connection is habitual for the physiologist. Stability is achieved through the auto-regulation of the internal environment of the organism. From the point of view of cybernetics this area of physiology offers a large number of instances of action in the organism of negative inverse connections; that is, such connections as those in whose presence changes and movements in the system (for instance the increase in the heart beat, the rise in temperature, etc.) are signalled to it and are accompanied by compensatory movements in the opposite direction. The precision with which stability is maintained of the pH of the blood, for example, well illustrates the importance of the negative inverse connection in the stabilization of the internal environment. A less obvious example of the inverse connection is that of the reflex activity of the organism under the influence of factors of the internal environment acting through distance receptors, *i.e.*, receptors such as the eye and ear, which are sensitive to stimuli having their origin at a distance. But in actuality it is scarcely possible that any reflex could function without the intervention and participation of an inverse connection.

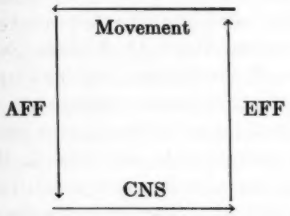
In an artificial mechanism the capacity to complete a task might theoretically be achieved by means of organizing and perfecting a whole complex of control devices. However, it would seem to be more advantageous, and more correct in terms of principle, deliberately to permit variability of the forces acting on the control centers, and to provide in this system for a negative connection of the inverse type: errors and deviations in the fulfillment of the task must be signalled to the system of control, which in response to such information would assure the compensation of such developing deviations. And if this inverse connection acts uninterruptedly, then the task can be fulfilled with any desired precision.

In nature, organisms may constantly anticipate accidental or unforseen circumstances. Therefore in living organisms the principles of the inverse connection have been used, it would seem, since life began. Gray Walter believes that the inverse connection played a leading role in the development and evolution of the nervous system. From this point of view, for example, the transfer of the higher nervous centers and the receptors into the cephalic portion of the organism could have progressive significance in evolution, because the organism, in moving about, is constantly signalled regarding environmental changes. In 1956, at the international conference on cybernetics in Namur, Belgium, there was a demonstration of the presence of mechanisms of the inverse connection in the amoeba.

Soviet scientists have frequently concerned themselves with the significance of the inverse connection for the organism. A. F. Samoylov expressed some ideas on the universal distribution in organisms of cyclic rhythms of stimulation (A. F. Samoylov, "Selected articles and speeches," Publications of the Academy of Sciences USSR, 1946). As Bernshteyn has shown, the coordinated movement activity of the living organisms and the unique character of their orientation in the environment cannot be explained without the concept of the inverse connection (N. A. Bernshteyn, "On the nature of movement," Medizdat, 1947). The uninterrupted inter-relationship between the receptor and the effector processes makes possible the fully controlled nature of movement in any changes of the environment; of composition of movement, of stress, of original posture, of original rate of movement of the effector organs, of displacement of the body, etc. In Bernshteyn's view, considered by the writer to be valid, the very abundance of "degrees of

freedom" (or possibilities of diversified reaction) which are allotted to the organism, in itself bears witness to the fact that at the basis of the control of reflex reactions nature has placed, not rigidly determined connections, but the possibility of the arrangement of reflexes and of their revision, depending upon circumstances in accordance with the correcting inflow of afferent impulses.

With very few changes, the scheme of action of the principle of inverse connection in the reflex activity of the organism, according to Bernshteyn, can be schematized as follows:



In this reflex chain, the afferent impulse which evokes stimulation is accompanied by activity in the central nervous system, which, through the efferent nerves, elicits in the effector a movement reaction. To this point the scheme does not differ from the usual scheme of the reflexes. Meanwhile, in the activity the new position of the body or of the organ is instantly evaluated by means of proprioception (in the broader sense of this word, that is, including the distance and vestibular signals as well). The resulting flow of afferentation carries in itself information about the environment, but also an uninterrupted account of the course of the reaction (and of the movements in the organism causing it, and of the factors acting upon it). The central nervous system uninterruptedly receives information on the degree of fulfillment or of unfulfillment of the task, as well as on the condition and preparedness of the receptors. This information is used for the inter-position of necessary corrections in the composition of the reaction or of its separate components. The inverse connection ensures continuity and the adaptive character of the reaction, which may be very complicated, changing its composition as circumstances vary.

From this one may understand why the complete or partial disruption of the inverse connection, that is, the loss of the information from the effectors due to damage of the receptors or of the afferent nervous connections (for example in the spinal cord or in the cerebellum), leads to serious aberrancies of movement coordination. At the same time the damage of the efferent paths has consequences; not primarily in discoordination of movement, but in the collapse of segments of the chains of reaction, with the preservation of its composition in the undamaged segments.

In proportion to the complexity of the movement reactions, the visual analyzers play an ever greater inverse connection role in the fulfillment of these reactions. As an example, one may think of the basic visual reflexes, in which fixation of the eyes on the object, optimal conditions of illumination, sharpness of focus on the retinal membrane, superimposition of the two images, and so on, are achieved according to the principle of the inverse negative connection. Deviation of the eyes to either side of the optimal position, or a deviation of the size of the pupils from the appropriate optimal illumination, serves as a stimulus for the movement of the eyes or for a change in the diameter of the pupils in the opposite direction.

The eye constantly sends to the central nervous system a double set of afferent impulses: impulses allowing an evaluation of the rotation of the visual axes, and impulses signaling movements in the visual field. It is the combination of these two moments in information received by the central nervous system at the time of the visual act, which assures the ability of the organism correctly to orient itself in the external en-

vironment with the aid of vision, independently of produced movements.

The principle of the inverse connection also playes an important role in more complex reflex acts. A man extends his hand and picks up a pencil: the reflex is accomplished correctly due to complicated inverse connections, acting simultaneously through the visual and movement analyzers. A man stands: this also is the result of complex control accomplished by inverse connections of the vestibular apparatus, of the movement and visual analyzers, etc. A not less important role is played by the principle of the inverse connection in behavior at even higher levels. Thus, adaptation to the environment requires a constant influx of inverse information.

The inadequacy of the generally accepted scheme of reflexes for the explanation of the adaptive character of conduct of the animal is emphasized also by P. K. Anokhin ("Problems of psychology," 1955). However, in this proposed scheme, the organism takes account of the information from the effectors not uninterruptedly and in the course of the reaction (in which the very essence of the principle of the inverse connection consists), but at the end, upon completion of the task. Hence, the adaptive composition of the reflex acts remains, in essence, unexplained. If, for example, a dog's brain receives signals about the fact that he has reached a feeding place, these signals do not carry information during his approach to the feeding place. Therefore, such a type of inverse connection cannot provide an explanation of the adaptive character of the composition of reflex acts. It is necessary to emphasize that this is not explained by the phenomenon of repetition, upon which the classical theory of conditioned reflexes depends, for as the result of repetition the stimulus must follow definite nerve pathways. In actuality, however, the dog's approach can be somewhat different each time and the movement composition be changed depending upon new, changed conditions. For instance, if an obstacle appears, then the dog goes around it and then to the place of feeding; possibly he runs around from the other side, and if the feeding trough is moved, then he runs after it, etc.

The theory of the inverse connection permits us to understand how the higher nervous activity of the animal can at one and the same time be both adaptive and variable, changing in accordance with new circumstances. The equilibration of the animal with his environment proceeds uninterruptedly, and in particular, according to the course of reflex reactions. The organism here is an integral part of a changing environment, and in this sense the theory of inverse connections in biology continues the line of Pavlovian "nervism" and the theory of conditioned reflexes.

The mechanism of the inverse connection plays a definite role not only in the conditioned, but also in the unconditioned reflexes, which also proceed under the control of information received from the receptors according to the course of performance of the reflex. Hence, the theory of the inverse connection throws new light on the problems of the variability of unconditioned reflexes in ontogenesis and phylogenesis, on the overgrowth of them in time connections, and so forth. This theory is of special significance with respect to problems of ecology and of instincts. In each reflex both positive and negative inverse connections occur in complicated sets, the more precise determination of which is one of the urgent tasks of physiology.

Problems of the formation of conditioned reflexes cannot be separated from problems of memory. On the basis of memory, time connections can also be formed in machines. In the model constructed by Gray Walter (the artificial turtle Cora; a conditioned reflex analogy), the combination of sound and light are "kept in mind," since they lead each time to a prolonged elevation of electrical potential in a certain unit. When

this potential, as the result of repeated combinations, attains a definite level, the time connection is closed and long-sustained electrical waves elevate the sensitivity of the output (effector) circuit to the "conditioned" sound. Now the sound elicits a reaction (movement of the "turtle"), which earlier could only be elicited by light (W. G. Walter, *The Living Brain.* Norton and Co., New York, 1953).

In K. Shennon's artificial mouse, time connections are also formed, and the movements of the mouse in a labyrinth to artificial food is each time quickened and made more precise. The remembering device of the mouse maintains and evaluates acquired experience, elevating the sensitivity to the elements in the environment which in previous experiences have led to the best results. Shennon's chess player, of recent construction, adapts its play to the tactics of the opponent. The memory devices mathematically work through the results of previous moves and "provide for" the probable future strategy of the opponent.

At present the mechanism of "memory" and the principle of the time connection are widely employed in various automatic industrial machines.

Cybernetics distinguishes two categories of "remembering" devices; transient and those of prolonged action. For mathematics, the transient, short-time, or "operative" memory is the "memory" of intermediate results upon the completion of complex calculating operations. In the organism, operative memory consists of the traces of effects of stimulation and inhibition. N. Wiener does not include in the category of memory the phenomenon of aftereffects in the nerve (the potential of aftereffect, the reflex phase, accommodation and so forth), although an analysis of these phenomena in the present connection would be of considerable interest. Operative memory in the organism assures the storing, tracing and chain character of the reaction. It is of interest that P. S. Kupalov and B. N. Lukov related the conception of the mechanism of trace reflexes to the idea of the circulation of impulses of stimulation through neuronal circuits during the period of inactive break in the reflex (P. S. Kupalov and B. N. Lukov, "Archives of the biological sciences," 1933, **33,** Nos. 5–6). This conception, to some degree, applies to all reflexes, insofar as they occur in time, with each link of the reflex chain influencing all subsequent links.

In accordance with one of the model variants of "memory," a certain impulse or set of impulses can circulate for a long time through closed circuits of neurons. In order for the stimulus to leave the neuron circuit, an additional signal is required in the direction of the effector neuron, which has a doubled threshold of stimulation and receives impulses continuously from the neuronal circuit.

Other types of connections are also possible for operative memory. One of them is used, as described, in the model of Gray Walter. Operative memory may not be confined to neuronal processes. Hormonal processes may also be involved, wherein the essential role may be played by photochemical, electro-chemical and other specific mechanisms.

Prolonged memory in the organism assures the strengthening of the conditioned connections. It is represented as a prolonged shift in the magnitude of the thresholds of transmission of stimuli across the nerve synapses, according to definite arcs. One could set up an analogy between this memory and the "external memory" of the electronic computing machine.

Problems of organic memory relate to two branches of cybernetics; the theory of information and the transmission of information through channels of connections.

Information theory is a strict mathematical discipline, in which statistical methods are applied to the problems of transmission and reception of signals in the presence of impediments. Important contributions to the mathematical theory of in-

formation have been made by the Soviet scientists A. Ya. Khinchin, A. N. Kolmogorov, V. A. Kotel'nikov, among others. Later, the methods of general analysis of the transmission of information were given a detailed foundation by K. Ya. Shennon ("The theory of transmission of electrical signals in the presence of hindrance," Publishing House for Foreign Literature, 1953), and certain problems of information theory were developed by N. Wiener.

In biology the theory of information suggests new paths for the qualitative and quantitative study of the activity of the system of analyzers (*i.e.*, of sense receptors).

The great service of cybernetics in this branch is the introduction of precision into the concept of information. From the point of view of Pavlovian physiology of conditioned reflexes, information carries with it new signals which are not indifferent and hence are accompanied by orientational reflexes. But what is actually the difference between a new stimulus and an indifferent one? It is understood that it cannot be held simply that all indifferent stimuli are extinguished stimuli, for in this case they would constantly appear upon the application of inhibiting action to the organism.

The theory of information suggests quantitative criteria of the degree of information or of innovation contained in a given communication, that is, in a definite set of received signals. The concept of the degree of information is connected with the probability of the appearance of signals. In a certain imprecise form, one can, according to the theory of information, consider that the communication will contain as much more information or innovation, as the number of improbable signals contained therein. The most probable, naturally, are the signals which are reproduced stereotypically in every communication. These are the indifferent signals. In the most clear-cut form, the orientational reflex arises in the presence of improbable signals, that is, in a communication containing such signals as, in the given circumstance, are rarely encountered. These signals bear information and thus attract interest.

The degree of information received by the central nervous system is not diminished if there is no transmission along the nerve of stereotypically repeated signals, which are therefore devoid of information. Just as in machines, there is no reason to load the channels of connection with such signals. It is expedient for the communications to be "codified," that is, to be transciphered, so that the channels of connection or of the nerve are capable of transmitting the maximal volume of information.

In the organism, the task of clearing stereotyped signals is taken care of by mechanisms of adaptation. The stereotype also takes place within the limits of a single stimulus, if it is "monotonous" in character. Therefore, the effects of the inclusion or exclusion of a stimulus or of a change in its structure are transmitted with special intensity through the nervous system, while the monotonous stimuli are accompanied by the phenomenon of adaptation.

The theory of information underlines the progressive significance of the binary system of computation—in the language of biology it is the law of "all or nothing," according to which nervous stimulation is disseminated in the organism essentially by means of transmission along nerve tissues of impulses of like magnitude. The binary system, employing only two signs, allows the quicker performance of complex computing operations, in comparison with other systems of calculation. All intermediate operations in this case are maximally simplified, since the individual signals bear the simplest standard form of the isolated impulses, which in turn leads to a maximal simplification of the channels of connection and of the transmitting devices.

The computing operations of electronic machines are outwardly analogous to certain forms of activity of the nervous system.

There it is the relay, and here the synapse, which assures the progress of the impulses along the channel of connection in different combinations. The transmission of variations in the intensity of stimuli is accomplished by means of changes in the frequency of the transmitted impulses, and not in their intensity. Information theory shows that the law of "all or nothing" and the modulation of the frequency of impulses in the nerve fiber is an optimal solution from the point of view of the universal problems of the transmission of information along channels of connection. The growth of the living world has proceeded along the path of securing, in the organism, this most perfected method of transmitting information.

In biology, so-called "codification" is to be understood as the transciphering of the impulses of stimulation which assures the reception of the signals and their tracing along the reflex arcs with a minimum of information loss. The structures of the codifying neuronal fields for various functional systems of the organism must be different, because in each instance the receptor fields and the effector organs have different characteristics in relation to the transmission and the reception of information.

Speaking of operative memory and adaptation, we have already noted the different methods of codification in the organism. The morphological basis for codification is the peculiarity of the structure of the neuron-field receptors or of the synapses. It is important, in principle, that the results of codification be not only displacements in the time parameters of stimulation, but also effects of the spatial distribution of the stimulus. Changes in the temporal relations are thus capable of being transformed into changes in the localization of the stimulus.

The theory of codifying systems poses new problems to the neurohistologist and the morphologist. One of these tasks, subjected in a certain degree to mathematical analysis in the literature, is devoted to questions of cybernetics connected with the problem of the nerve "funnel," *i.e.*, the transmission of the stimulus into a diminishing number of neurons. Without becoming involved in the mathematical details, we shall present here two results of the calculations of A. Rapoport ("Bulletin of Mathematical Biophysics," 1955, **17**, No. 1), bearing on the visual analyzer.

Information accumulated in 130 million receptors of the retina must, with minimal loss, be transmitted along 800 thousand fibers of the optic nerve. In certain circumstances it is possible to compute the number of neurons which, with optimal transmission of information, ought to be found in connection with definite channels of the contracting "funnel." As shown in Figure 1, taken from the work of A. Rapoport but slightly modified in accordance with his conclusions, in the transmission of information from the retina with N neurons into a two-neuron chain, one third of the neurons N must have two connections and two thirds must have one.

Calculations also reveal another important consideration. Upon the transmission from a larger number of elements to a smaller, the loss of information will be reduced to a minimum only when the codifying network contains a "memory" device.

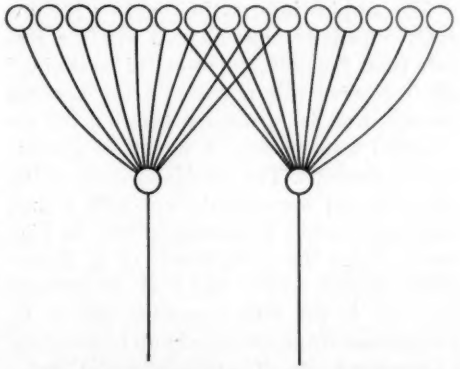


FIG. 1. Optimal arrangement of connection of a number of neurons with two neurons, in accordance with information theory (after Rapoport).

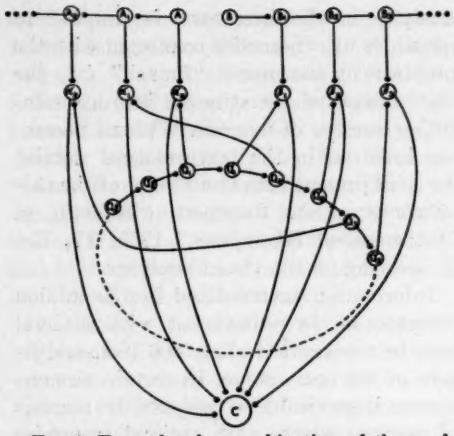


FIG. 2. Example of a combination of the work of the remembering and scanning devices according to information theory (after Rapoport).

This is explained by the fact that in the presence of the narrowing funnel the correct transmission of information essentially occurs at the expense of an increase in the duration of each communication. With this, the surplus, yet-untransmitted information accumulates intermittently and temporarily in the "memory" device. Thus, strictly speaking, in each reflex the elements of memory participate, for the signals accepted by the receptors and directed along the "funnel" must be subjected to codification in conformity with the morphological structure of the given analyzer system.

The surplus information held by the operative memory must be used within a certain time. Possibly, the so-called "scanning" devices serve this purpose in the nervous system, that is, structures assuring the sequential transmission of signals in a number of elements. The combined work of the memory and scanning devices with a time reading is shown in schematic form in Figure 2 (from the same article of A. Rapoport). In this scheme the chain of neurons $Q_1 \cdots Q'_i$ is the time computer and A, B, the neurons transmitting stimuli to the chain of neurons A_i, B_i. The threshold of stimulation of the neurons $alpha_i$ and $beta_i$ are elevated two-fold. As the unit of time, the interval of transmission of stimulus from Q_1 to Q_1' is taken, and so forth. Then it is apparent that $alpha_0$ is stimulated only at the moment t equals unity, and only in case A is stimulated at the moment t equals zero; $beta_1$ is stimulated at the moment t equals two, if B was stimulated at the moment t equals zero; $alpha_1$ is stimulated at the moment t equals three, if A is stimulated at the moment t equals unity; $beta_2$ is stimulated at the moment t equals four, if B was stimulated at the moment t equals unity. Insofar as $alpha_i$ and $beta_i$ transmit the stimulus to neuron C, then the informative communication is transmitted along a single chain, changes structure and is prolonged in time.

By computation, overloading with information takes place not only in definite peripheral organs or neuronal fields, but also in the brain. Therefore in cybernetics it is admitted that the neurons of the higher centers of the nervous system periodically are "scanned" by special reverberating systems of neurons. It was initially accepted that scanning in the human brain occurs in the rhythm of the alpha-waves, which would take on, in this way, a functional interpretation. However, this suggestion, despite a number of works conducted on the basis of it, has not yet been confirmed.

It is possible that scanning also has significance for the processes of generalization or inference. On this was based the scheme of the apparatus for reading lettered text by the blind, in accordance with the theory of W. S. McCulloch and W. Pitts ("Bulletin of Mathematical Biophysics," 1944, 5, No. 2). The unusual peculiarity of this system lies in the fact that it permits reading of letters independently of their size in the text. Whatever might be the size of a given letter, it is communicated by one and the same set of sounds—one and the same melody. It is noteworthy that the scheme of the scanning device of McCulloch and Pitts, according to the testimony of the well-known neurohistologist G. V. Bonin

(*Essay on the Cerebral Cortex*, Thomas, Springfield, Illinois, 1950), in its basic features compares with the structure of the fourth layer of the visual cortex.

In these conclusions information theory closely touches on the problems of the second signal system.

Applicable to the organism are not only the problems of codification of the efferent impulses, but also problems of the transciphering of the stimulus for its coordinated emergence into the effector pathway. On the plane of the second signal system this means that the theory of information has a direct relation not only to perception and its generalization, but also to verbalization of information processed in the brain. In man a highly complicated codification of a large amount of generalized information must occur in the nerve structures for transmission with minimal loss through the vocal organs in verbal communication.

With the evolution of the living world, methods of reception and transmission by the organism of information in maximal amounts have been undergoing improvement, as have its methods of generalization and further transmission. In this way the nervous system long ago defined all the robot mechanisms created by man.

The principle of control includes the conception which N. Wiener formulates thus, "Information is information and not energy."

The energy contained in informative communication is insignificant, and the transmission of it, upon the accomplishment of control and of connection, plays practically no role in respect to the quality and volume of information.

Control of machines or of organisms is the more complete the greater the volume of processed information, and the degree of information, as already indicated, is basically the improbability of a given communication. With the transmission of informative communication the distribution in time and space of the impulses of stimulation in the nervous system takes on structural forms which are less probable the more valuable the information received. In this sense the information acts as a counterweight to the second law of thermodynamics, which requires that Nature in all its manifestations strives toward the most probable state, that is, to an absence of organized structures.

Thus information must be regarded not as the source of energy but as the organizing principle, and it is precisely this which accounts for its great value in living nature and in technology. The value of information is so great that neither man nor Nature regrets the expenditure of energy in the securing of its reception and transmission. To transmit information without hindrance is possible only in those cases when the receiving and transmitting devices are in a stable state. Hence the flow of metabolic processes at a definite level and the stability of the internal environment are necessary conditions for optimal functioning of the systems of control and connection both in the machine and in the organism. Cybernetics, then, and at a somewhat new level, returns to the well-known position of Claude Bernard, who remarked that, "the stability of the internal milieu is a condition of the free life."

In an absolute sense, of course, informative processes cannot be separated fully from the energy processes. In this connection N. Wiener does not exclude the possibility of an influence of the structure of informative communications on the growth and morphology of the structure of tissues and organs. The intermediaries, it seems, can be specific enzymes to which, thereby, may hypothetically be attributed the role of the so-called "demons of Maxwell," that is, of formations employing the signals from the environment for the organization of structures in counterpoise to the second law of thermodynamics.

In this hypothesis cybernetics begins to touch upon the problems of morphogenesis.

In discovering definite analogies and points of similarity between the nervous system and complicated robots, cybernetics, nonetheless, does not offer physiology ready solutions to the complex problems of nervous regulation and of nervous activity. Even more, cybernetics refrains from equating the highest and most complex organ in Nature—the human brain—with automatic computation machines.

Discovering in technology the great importance of the principles of information and of the inverse connection, cybernetics has revealed that Nature constantly employs them in the living organism and that these mechanisms, hence, actually have general significance.

Along with certain new points of view, neurophysiology, thanks to its contacts with cybernetics, has also acquired new quantitative methods of work from the arsenal of mathematics and of biophysics. This development is consonant with the ideas of Pavlov, who anticipated and welcomed the contact of physiology with the strict scientific disciplines.

In conclusion it is necessary to remark that, for contemporary brain physiology, there have appeared a number of important ideas, also relating to the principles of control, but going beyond the bounds of ideas used at the present time in cybernetics. One may expect that further contact of cybernetics and the physiology of the nervous system will give new, mutually-fruitful results which will exercise a powerful influence on the growth of the science of control.

BRIEF COMMUNICATION

A NOTE ON THE THEORY OF HOMOSEXUALITY: THE BODY AS OBJECT

ROBERT SEIDENBERG, M.D.[1]

Early in the evolution of psychoanalytic metapsychology, Freud (2) wrote of two classes of instincts, sexual and ego. One premise for this dichotomy was the impression that the sexual instincts alone sought out or needed external objects for their gratification. The division on this premise was no longer tenable when it was realized that sexual instincts were often self or body-directed and gratified. As such they could no longer be differentiated from ego instincts. Largely for this reason, this division between instincts was abandoned.

To explain the phenomena related to the self and external objects, new emphasis was placed on "energy" and its direction. (The concept of life and death instincts is not pertinent here). Libido, then, could be directed toward an external object (object libido) for gratification or to oneself (narcissistic libido). The latter event was thought of as a regressive phenomenon. Here, the objects are dichotomous and stationary, in contrast to the energy of the drives which is unitary and may shift from the external object to the self and vice versa.

Fairbairn (1) abandoned the concept of libido shifts and emphasized degrees and qualities of object relatedness. To some, this approach seemed more consonant with clinical data, especially of psychotic conditions where external personal objects appear to be entirely lost. Here the "hold" on objects appears to preempt drive gratification. In this construction, need and object are inextricable.

Recently, Szasz (4) explained bodily feelings in schizophrenia as a consequence of loss of personal objects. He stated that following the loss of object (reacted to with a mourning reaction or a paranoid mechanism), the feelings and fantasies may be transferred from the lost object to the body. The body then becomes the only remaining object, explaining the hypochondriacal preoccupations and dread that it too may be lost. Rather than speaking in terms of shifts (regressive) from object to narcissistic libido, Szasz adroitly used the concept of transference to explain the schizophrenic's new relatedness to his own body.

The point of departure of the present communication is Szasz's (5) observation that "homosexuality" represents the individual's predominant interest in his own body rather than in another of the same sex as might appear to the observer. Material from the analysis of a "borderline" case exquisitely demonstrated this phenomenon.

A thirty-two year old research scientist came for treatment on his arrival in this city. He had been in therapy for about two years at a previous residence. He had many hypochondriacal symptoms, the most distressing of which was his fear of eating solid foods. He felt that the musculature of his swallowing apparatus would fail him and he would die of suffocation by aspirating the food into his lungs. This caused him to subsist on liquids which he felt would pass through with less danger. With this he experienced ptyalism. At times he was unable to swallow the saliva and experienced extreme discomfort.

He sought treatment initially because of homosexual feelings although there was

[1] Department of Psychiatry, Upstate Medical Center at Syracuse, State University of New York.

never any overt activity. He disclaimed interest in women, fearing the occasion when he would have to kiss one. Masturbation was frequent. He freely recalled his fears that his mother might become impregnated by his semen around the toilet bowl. He had an older brother toward whom he had a passive relationship.

In his work, he had bitter resentment toward his department head for inequities, real and fancied. He was fretful because he thought he was not progressing rapidly enough and would brood if others in his field were ahead of him. There was never enough time for his research, teaching and other intellectual pursuits. He reluctantly accepted social engagements. They interfered with his work but he condescended to attend because "one must socialize."

The pertinent material came in the sixth month of the second course of therapy. He had been to a scientific meeting. He said he felt depressed because he had heard of work being done in another area that was unfamiliar to him. He blamed the department head for not providing the proper periodicals. Two days after his return from the meeting he related the following incident.

"Yesterday afternoon as I was walking downtown I saw a pair of swimming trunks in a store window. Looking at them on the manikin, I became sexually excited. I went in and bought a pair. They were especially short. When I got home, I put them on and looked at myself in the mirror. I saw immediately that I could not wear them like that on the beach because the outline of my genitals was clearly visible. One would have to wear a jock-strap. In watching myself and nothing else I became sexually excited and masturbated."

The night following the above incident, the patient had the following dream.

"I was in my bedroom. My brother came in. I was worried lest he find my swimming trunks. I wondered where I had put them and went into the next room to look—but without success. When I returned to the bedroom, there was my brother on the bed, wearing my trunks. I was embarrassed and wondered what he was thinking. I noticed that his genitals did not show through the trunks."

In his associations to this dream he stated, "I suppose this dream points up my basic homosexual feelings which have been plaguing me since I can remember. The idea of a man's body is very exciting to me but I don't recall such feelings about my brother. I suppose my interest in the manikin with the short trunks was due to my frustrated homosexuality. I seemed to be impressed with the fact that my brother's genitals did not show through." On further urging, he added, "I look at my body a great deal. Having a thin waist and small buttocks excites me. I admire thin qualities in other men too. I don't find stout men sexually attractive. Thinness and leanness—principally around the buttocks and the back of my thighs. Keeping on my diet has been satisfying in that way. Before I became thinner, my lower abdomen sagged; this was distasteful. So the loss of weight due to my eating difficulty served the purpose of making my body more agreeable to me—more sexually interesting."

DISCUSSION

It appears from the masturbatory and and dream sequence that this patient was relating to his body as an object. Through the mechanism of the dream, he attempted to transfer this interest to the body of his brother. Apparently the prior preoccupation and gratification with his own body caused intense misgivings in him and led to an attempted externalization. The dream helped him feel that there was an object outside of himself. As such, the dream can be viewed as a wish for restoration of objects which were disappearing and were threatening him with complete object loss. In terms of body image the fact that he

could see his own genitals in contrast to the dream in which his brother's were not seen, is also a reassurance for him that this part of his own body is still intact. The masturbatory activity itself could also serve this purpose. His fantasy might be paraphrased as follows, "I am safe, I can see my genitals through the trunks; it is my brother's which aren't seen and may not be there at all."

The bathing trunks serve as a transference vehicle from the masturbatory episode to the dream. They likewise represent the area of the body which is in greatest danger of being lost to him; the area most likely to succumb to the forces within him which are working to exclude his masculinity from his body image. His defensively hypercathected intellectuality was usurping even his interest in his body and was threatening to leave him with "thinking" as his only tolerable object.

In traditional psychoanalytic terms, the foregoing process and phenomenon is akin to the narcissistic object choice as described by Freud (3). Here the object chosen is: a) what he is himself b) what he once was c) what he would like to be d) someone who was once part of him.

SUMMARY

A case is presented where the body of the individual is the last remaining object. A sequence was revealed that demonstrates an attempt at restoration of an external object via homosexual fantasies. Here the homosexual interest was thought to be derived from a transference from interest in one's own body to that of another.

REFERENCES

1. Fairbairn, R. *An Object Relations Theory of the Personality*. Basic Books, New York, 1954.
2. Freud, S. *Civilization and Its Discontents*, Ch. VI. The Hogarth Press, London, 1951.
3. Freud, S. *On Narcissism; An Introduction*. In *Collected Papers*, Vol. 4. The Hogarth Press, London, 1925.
4. Szasz, T. *The Psychology of Bodily Feelings in Schizophrenia*. Psychomantic Medicine, XIX, 19: 11, 1957.
5. Szasz, T. *Pain and Pleasure, A Study of Bodily Feelings*, p. 141. Basic Books, New York, 1957.

BRIEF COMMUNICATION

NOTE ON THE CONCEPT OF "SPLIT SOCIAL FIELD" AS A DETERMINANT OF SCHIZOPHRENIC EXCITEMENT

EUGENE B. BRODY, M.D.[1]

Stanton and Schwartz, building upon the ideas of Harry Stack Sullivan, have proposed that a "symptom" is stable only because it is part of a stable social equilibrium. The matrix of interpersonal relations in which symptoms occur is conceived as a social field "in which the patient's activities are responses and adaptations. . ." (8). This social field may be "split" by covert conflict between staff members upon whom a patient is dependent. In that event, the schizophrenic patient's excitement or dissociation is considered to be "a reflection of, and a mode of participation in" the split social field (7).

A number of other, more familiar psychoanalytic concepts may be useful in understanding schizophrenic excitement in this type of context. Freud's suggestion (1) that aggressive behavior might be a response to deprivation was elaborated upon in the frustration-aggression hypothesis of Miller, *et al.* (4). Hartmann, Kris and Loewenstein (2), discussing the importance of childhood deprivation for the differentiation between self and outside world, believed that the tension accompanying each transition from indulgence to deprivation is tolerable only when the child is secure and confident that indulgence will again follow. Considering aggression as a basic drive, they hypothesized (3) that unfulfilled libidinal expectations (*i.e.* deprivation of libidinal discharge) could result in a substitutive aggressive discharge. B. Rank (6) was impressed by the narcissism of the mothers of those children, who, unable to tolerate deprivation, exhibited frequent aggressive excitement. She concluded that in such children the ego remains incapable of organizing and controlling aggressive and libidinal drives, and that their tantrum-like reactions often seem to be panic, rather than rage, as if in response to threats of annihilation.

These concepts may throw additional light on aggressive, excited behavior. Rackow (5), for example, studied the factors related to aggressive acts during a two year period in a large psychiatric hospital. He reported peaks in the incidence of aggressive behavior before meals, at bedtime, in the toilet areas, while waiting for required and disliked activities, and following interference with an ongoing activity by another patient.

It is suggested that these incidents occurred with greatest frequency when accustomed patterns of need-gratifying behavior were interrupted (interference with libidinal discharge), and when certain expectations and needs were aroused, but not gratified. That is, they occurred in response to deprivation which emphasized the difference between self and object world, and the threatening, unreliable nature of the latter. Aggressive behavior in this instance may be considered both as a substitutive energy discharge in the face of libidinal deprivation, and as an expression of panic occasioned by interference with the artificially ordered world constructed by the schizophrenic to allay his anxiety.

The writer has observed a number of excited episodes in chronically hospitalized schizophrenic men whose symptoms could be partially understood as reflecting a form of participation in a relatively stable social situation. In several instances aggressive ex-

[1] The Psychiatric Institute, University of Maryland School of Medicine, Baltimore.

citement occurred in a context of disagreement between staff members upon whom the patient was dependent. Of particular interest was the observation that the patient was usually more dependent upon one staff member, a "key person," than upon the other, and that this staff member *himself* was in conflict about the issues pertaining to the patient. The conflict within the "key person" could be schematized as one involving wishes to gratify the patient's dependent needs versus wishes to be free of the relationship. The key person's conflict was inevitably reflected in a tendency to withdraw his libidinal investment in the patient.

These observations suggest that when a split occurs in the social field the patient is deprived of his usual libidinal gratification. This deprivation is not the result of the split in the social field, but of a change in his relationship with the key person, who is a major constituent of the social field. It may be that the vicissitudes of the patient's relationship with this person, who is his primary source of libidinal supplies, are the most important determinants of his stability or excitement.

A closely related hypothesis is that the executive and integrative ego functions of the chronic schizophrenic patient, existing in a limited hospital environment, fluctuate in scope and effectiveness with the relative intensity of gratification and deprivation from the person upon whom he is most dependent for libidinal supplies. This "key person," is, as were the mothers of the children described by Rank, the most important constituent of the patient's object world. When his own gratification is threatened (either because of conflict with some other staff person or for other reasons) he, without being aware of it, withdraws from the patient some of the libidinal supplies upon which the latter has become dependent. This constitutes not only a deprivation, but a major disruption in the patient's most important object-relationship. This is the relationship which we consider as the patient's bridge to the outside world and reality. The disruption of this relationship may be the most significant factor in the production of the kind of schizophrenic excitement with which we are concerned.

We may then suggest that the crucial "split" in the social field of the schizophrenic patient occurs *within* the key person who, by virtue of his particular interest, has become a source of libidinal supplies, an object of libidinal discharge, and possibly an object of identification for the patient. This is, in fact, suggested in Stanton and Schwartz's later consideration of the situation (9). Discussing a patient's excitement they state, "the two most immediately important persons in his life were, so to speak, pulling him in opposite directions, and each showed, while he continued doing this, that he, himself, was of a *divided mind*" (italics added).

REFERENCES

1. Freud, S. *A General Introduction to Psychoanalysis.* Garden City Books, Garden City, N. Y., 1935.
2. Hartmann, H., Kris, E. and Loewenstein, R. Comments on the formation of psychic structure. Psychoanalytic Study of the Child, **2:** 11–38, 1946.
3. Hartmann, H., Kris, E. and Loewenstein, R. Notes on the theory of aggression. Psychoanalytic Study of the Child, **3/4:** 9–36, 1949.
4. Miller, N., Sears, R., Mowrer, H., Doob, L. and Dollard, J. The frustration-aggression hypothesis. Psychol. Rev., **48:** 337–342, 1941.
5. Rackow, L. Aggressive behavior in a mental hospital. Presented at the Annual Meeting of the American Psychiatric Association, Atlantic City, N. J., May, 1955.
6. Rank, B. Aggression. Psychoanalytic Study of the Child, **3/4:** 43–48, 1959.
7. Stanton, A. H. and Schwartz, M. S. Observations on dissociation as social participation. Psychiatry, **12:** 339–354, 1949.
8. Stanton, A. H. and Schwartz, M. S. A social psychological study of incontinence. Psychiatry, **13:** 399–416, 1950.
9. Stanton, A. and Schwartz, M. *The Mental Hospital,* p. 362. Basic Books, New York, 1954.

BRIEF COMMUNICATION

POSITIVE GLYOXYLIC ACID REACTION IN THE URINE OF MENTAL PATIENTS

M. VELEK, M.D.[1] AND M. VALENTOVA, M.D.[1]

Recently Riegelhaupt (7) reported that the glyoxylic acid reaction, which is specific for materials of the tryptophan group, is positive in the urine of schizophrenics. In 65 urines of schizophrenic patients, Riegelhaupt found 51 positive with the glyoxylic acid test, six doubtful and eight negative; while in a group of 55 urines from non-schizophrenic patients, 40 were negative, seven doubtful and only eight positive. Twelve control urines from healthy persons were negative. The reaction used by Riegelhaupt is essentially analogous to the Adamkiewicz reaction which depends upon the trace of glyoxylic acid in the acetic acid used in the reaction (3). On the other hand, the reaction of Pospisil (6) is known to be positive, in the presence of aromatic compounds of the tryptophan group, in the filtrate of deproteinized serum of acute uremic cases; this reaction uses the underlayer of sulphuric acid without the presence of acetic acid. It appears therefore that neither acetic acid with a trace of glyoxylic acid nor glyoxylic acid is invariably necessary for the positive Adamkiewicz reaction.

We have investigated the Adamkiewicz reaction using acetic acid without any glyoxylic acid contaminant as well as acetic acid containing a trace of glyoxylic acid, and we have investigated the Riegelhaupt reaction with glyoxylic acid as well as the similar reaction using only sulphuric acid.

We carried out the Adamkiewicz reaction using acetic acid with a trace of glyoxylic acid (in our case concentrated acetic acid P.A., Lachema) and, in parallel experiments, the reaction with acetic acid without a trace of glyoxylic acid. In both cases we mixed 1 cc. of urine with 1 cc. of acetic acid and underlayed 1 cc. of concentrated sulphuric acid. We called the reaction positive if a brown-violet ring formed in a short period of standing.

We prepared the acetic acid uncontaminated by glyoxylic acid by boiling under reflux for 1 hour 1 liter of acetic acid with 5 grams of potassium permanganate and then distilling the acetic acid over in a completely glass apparatus. We used the Brice (1) test for glyoxylic acid as modified by Langan (4): In a test tube we underlayed 0.95 ml. of the distilled acetic acid with 0.65 ml. of concentrated sulphuric acid. We cooled it and added 0.05 ml. of a 100 mg.% aqueous solution of tryptophan. We maintained the mixture in a water bath for 20 minutes at a temperature of 56° C. Pure acetic acid gave only a yellow or orange color.

We carried out the Riegelhaupt reaction by mixing an equal volume of 5% glyoxylic acid with the urine and underlaying it with 1 ml. of concentrated sulphuric acid. At the same time we ran a parallel experiment in which the urine alone was underlayed with 1 ml. of sulphuric acid.

All urines were collected in the morning before breakfast. Urines which were cloudy or contained sugar or protein were discarded. The diagnoses of the patients were known only after the urine analysis. We used only clear, unfiltered urines. One hundred and sixty-one urines were investigated: 55 urines from schizophrenics most of whom were chronic, 34 from chronic alcoholics, 28 from arteriosclerotic patients most of whom were demented, seven from oligophrenics,

[1] The Central Laboratory, The Psychiatric Institute, Praha 1, Kozi 10, Czechoslovakia.

six from cases of progressive paralysis, three from epileptics, four from paraphrenics, and 24 from other psychiatric patients (manic depressives, psychotics, Pick's disease, etc.). The group of 34 alcoholics we used for a normal group.

None of the investigated reactions proved specific for any diagnostic group. However, all four reactions were more positive in the schizophrenic group than in the normal group or in any other diagnostic group. All four investigated reactions gave similar results on the whole; therefore, glyoxylic acid is not an irreplaceable reagent for the tryptophan group.

Urines which were filtered over active charcoal gave negative reactions. Standard solutions at concentrations of 100 mg.%, 10 mg.% and 1 mg.% of the nitrogen compounds in urine which could contribute to the positive reaction were tested. These nitrogen compounds are uric acid, urea, creatinine and tryptophan. All four investigated reactions were positive only with tryptophan in concentrations of 100 mg.% and 10 mg.%. The other standard solutions gave completely negative results in these reactions. The tryptophan we used was pure so that we do not believe that either iron or copper is necessary for a positive reaction as has been suggested by Hawk (2).

Since tryptophan is not directly excreted in urine we believe that the positive reaction depends on metabolites of tryptophan: hydroxyindoleacetic acid, potassium indoxyl sulphate and perhaps also traces of kynurenic acid. The indoleacetic acids are urinary chromogens from which are formed the urinary dyes, urorosein and uroerythrin, which may be identical with one another. Since the reaction seemed to be more intensely positive the darker the urine, it would appear that the reaction may depend to some extent on the urinary chromogens. Furthermore urochrom gives a weak positive color reaction for tryptophan (5).

The excretion of chromogen is clearly variable from time to time in the same patient since the intensity of color obtained in the investigated reactions changes on repeated determination; this was also observed by Riegelhaupt (7) in nine out of 41 cases where a second urine sample was investigated.

An increased excretion of urinary chromogens is reported in the literature in chronic kidney disease, tuberculosis, liver disease and pellagra. Schizophrenics therefore constitute a further group of patients with increased excretion of urinary chromogens. Whether or not this elevated excretion of urinary chromogens is connected with an elevated excretion of toxic urine materials one cannot say. The reactions under in-

TABLE 1

Tabulation of Investigated Reactions Discussed in Text

	Reaction with Glyoxylic Acid		Reaction with Sulphuric Acid		Reaction with Acetic Acid with a Trace of Glyoxylic Acid		Reaction with Acetic Acid without a Trace of Glyoxylic Acid	
	−, ±	+, ++	−, ±	+, ++	−, ±	+, ++	−, ±	+, ++
Schizophrenics	18	37	23	32	15	40	18	37
Chronic Alcoholics	10	24	20	14	18	16	18	16
Arteriosclerotics	14	14	12	16	10	18	10	18
Oligophrenics	4	3	4	3	4	3	4	3
Progressive Paralysis	5	1	3	3	5	1	4	2
Epilepsy	3	0	3	0	3	0	3	0
Paraphrenics	2	2	1	3	2	2	1	3
Others	12	12	12	12	12	12	12	12

−, ± = Negative and slightly positive reaction.
+, ++ = Positive and very positive reaction.

vestigation are not specific for 5-hydroxytryptamine which is excreted in urine in the form of 5-hydroxyindoleacetic acid.

Whether the indole nucleus or the side chain takes part in these reactions remains to be discussed. Riegelhaupt (7) found peracetic acid (acetic acid peroxide) caused the disappearance of the color reaction with glyoxylic acid and explained this as an oxidation of the side chain. But since only the underlaying with sulphuric acid is necessary for the occurrence of a color reaction, one may believe that the basis for these reactions is the sulphonation of the tryptophan nucleus.

Table 1 presents the results of the investigated reactions in the various diagnostic groups.

This study indicates that the investigated reactions are not specific for any diagnostic group: only in the schizophrenic group is there indirect evidence of an increased excretion of urinary chromogens. This, however, is not proof of the excretion of a toxic material in the urine of schizophrenics.

REFERENCES

1. Brice, A. T., Jr. Tryptophane content of blood serum; new technic. J. Lab. & Clin. Med., **19**: 1113–1115, 1934.
2. Hawk, P. B., Oser, B. L. and Summerson, W. H. *Practical Physiological Chemistry.* London, 1949.
3. Kutacek, M. *Die Ubungen aus Biochemie,* p. 64. Pedag. Verlag, Prague, 1956.
4. Langan, T. A., Durrum, E. L. and Jencks, W. P. Paper electrophoresis as quantitative method: measure of alpha and beta lipoprotein cholesterol. J. Clin. Invest., **34**: 1427–1436, 1955.
5. Nosek, J. *Neue Grundlage der Biochemie,* p. 193. Brun, 1956.
6. Pospisil, V. Cas. lekaru ces., **81**: 844, 1942.
7. Riegelhaupt, L. M. Investigations on the glyoxylic acid reactions on urines from schizophrenic and other psychotic patients. J. Nerv. & Ment. Dis., **123**: 383–385, 1956.

BRIEF COMMUNICATION

AN UNUSUAL COMPLICATION OF ELECTRO-SHOCK THERAPY: MASSIVE RETROPERITONEAL HEMORRHAGE IN A PATIENT WITH BLEEDING TENDENCY

FRED LEVENTHAL, M.D.[1] AND HOUSTON BRUMMIT, M.D.

Over the past twenty-five years a large number of complications and contraindications to electroconvulsive therapy have been reported. These include reversible and irreversible changes in neurons, gliosis, memory loss, electroencephalographic changes, cardiac arrest and arrhythmias, status epilepticus, manic delerium, autonomic disorders, vertebral and other fractures, aggravation of hypertension, congestive failure, aspiration pneumonia and lung abscess (8). This paper records an unusual and rare complication of electroconvulsive therapy.

CASE REPORT

The patient was a 33 year old white unmarried male who was said to have "pseudohemophilia" from childhood. At age seven a tooth extraction led to severe bleeding which required hospitalization. Many medications including snake venom were tried until bleeding stopped, and multiple transfusions were needed.

In January of 1943 the patient was drafted into the Army, and in August, 1943 a tooth was extracted over his objections. Again, severe bleeding resulted and cold compresses, pressure packs, sutures and multiple blood transfusions were tried. The patient was sent to a large Army hospital where a diagnosis of "pseudo-hemophilia" was made.

Four years later, in January of 1957, the patient was admitted to the Brooklyn Veterans Administration Hospital with a diagnosis of Schizophrenia, Undifferentiated type. He was observed to be preoccupied, manneristic and hallucinating. He was unable to continue his work as a draftsman. His family reported that they had always tended to overprotect the patient because of his bleeding tendency. They remarked that he was a withdrawn, asocial individual. Physical examination of the patient at this time was unremarkable.

The Psychiatric Service felt that electroconvulsive therapy would be the treatment of choice, but that the history of a bleeding tendency made it desirable to use other therapies first. Chlorpromazine, Reserpine and Meprobamate were of no value. An investigation of the blood tended to show that the patient's bleeding was due to a circulating anti-coagulant. When the patient lapsed into a catatonic state it was felt that electroconvulsive therapy should be tried.

In August, five standard electroconvulsive treatments were given within a period of nine days. On the day following the fifth ECT, the patient complained of pain in the right buttock and right costovertebral angle. Two days later the patient developed splinting of the right abdomen with continued pain in the right buttock and flank. The hemoglobin before shock treatment was 13.8 gms/100 cc and one week after shock treatment was 10.4 gms/100 cc. Platelets were normal; there was no hematuria. The patient's symptoms and signs gradually disappeared.

[1] 950 Westbury Road, Westbury, New York. (Formerly, Department of Radiology, Brooklyn Veterans Administration Hospital, Brookly, N. Y.)

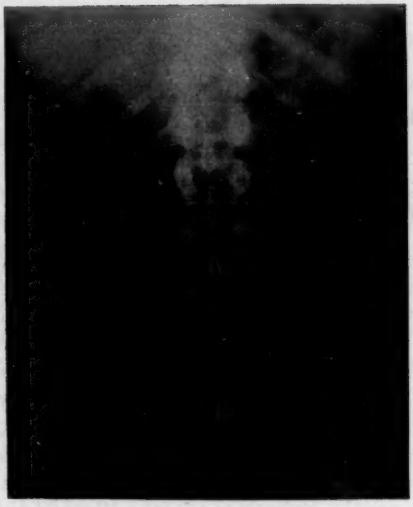


FIG. 1. Kidneys in normal position: right psoas muscle evident.

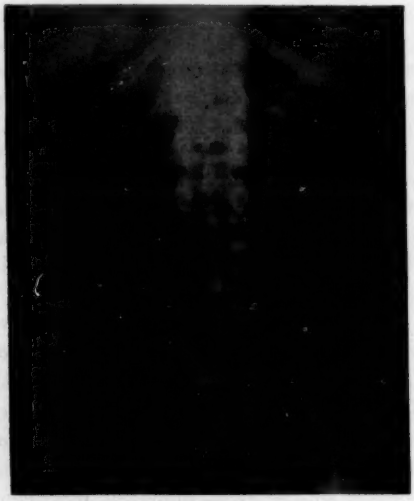


FIG. 2. Elevation and outward rotation of right kidney: right psoas muscle obscured.

X-RAY

On August 29, 1957, twelve days after the onset of right flank pain, radiographic examination of the abdomen (Figure 2) showed the right kidney to be elevated with its lower pole rotated outward. The psoas muscle shadow on the same side was partially obscured. In light of the clinical history, the impression was that of retroperitoneal hemorrhage on the right. This was particularly striking when contrasted with a film of the abdomen dated January 8, 1957 (Figure 1), at which time the right kidney and psoas muscle shadow were entirely normal.

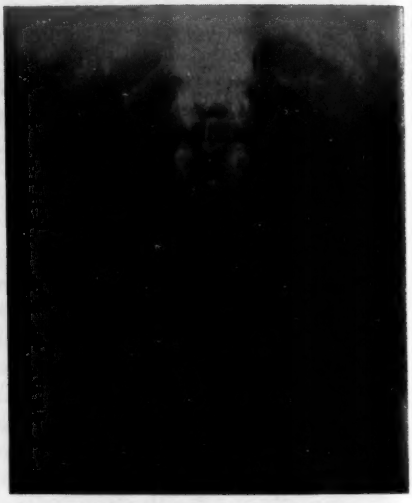


FIG. 3. Right kidney returned to normal position: right psoas muscle again evident.

On October 4, 1957 an intravenous pyelogram was performed (Figure 3). The right kidney had returned to its normal position and the psoas muscle again had become distinct.

DISCUSSION

The clinical course, laboratory and radiographic data are extremely convincing of massive retroperitoneal hemorrhage. The trauma of electroconvulsive therapy led to hemorrhage in this susceptible patient. However, other hemorrhagic and thrombotic phenomena have been reported in patients without bleeding tendencies. These include cerebrovascular hemorrhages, rare nosebleeds, occasional conjunctival hemorrhages, thrombophlebitis and coronary oc-

clusions (7). A rare case of hemorrhage from the lung in active tuberculosis has been recorded (9). Adrenal apoplexy (2) and thrombocytopenic purpura have also been recorded (6).

Study of the blood in human beings and guinea pigs following electrotrauma shows a slight transient increase in the prothrombin time and a slight increase in the heparin activity of the blood (4, 5). It has also been reported that electroconvulsive therapy may lead to transient (two to nineteen hours) increase in fibrinolytic activity which affects fibrin, though not fibrinogen (7). In a normal person this is probably not significant but in a susceptible person it may be enough to "tip the wagon."

The effect of electroconvulsive therapy on human capillaries has been described as showing thinning of capillaries, interruptions of the capillary stream with spasm and irregularities of the caliber of the capillaries with decrease in blood flow, which goes back to normal ten minutes after electroconvulsive therapy (5). There is also, during electroconvulsive therapy, a drop in circulating plasma volume, as shown by the Evans blue technique, averaging 18 per cent with a corresponding rise in the hematocrit (1). This tends to facilitate the development of thromboses.

While the effect of electroconvulsive therapy may predispose to hemorrhages and thromboses, these are probably relatively uncommon. To our knowledge this is the first example of massive retroperitoneal hemorrhage reported.

CONCLUSIONS

This paper records a most probable massive retroperitoneal hemorrhage in a young man following electroconvulsive therapy. The patient had a known bleeding tendency, probably due to a circulating anti-coagulant. The case demonstrates that bleeding tendencies are potent contraindications to electroconvulsive therapy.

REFERENCES

1. Altschule, M. D., Cram, J. E. and Tillotson, K. J. Hemoconcentration after electrically induced convulsions in man. Arch. Neurol. & Psychiat., **59**: 29–38, 1948.
2. Altschule, M. D. Adrenal apoplexy following electronarcosis therapy. New England J. Med., **253**: 191, 1955.
3. Calogera, E. Behavior of heparin activity in electrotrauma. Folia med., **34**: 517–521, 1951.
4. Calogera, E. Behavior of prothrombin time in electrotrauma. Folia med., **36**: 80–84, 1953.
5. Davis, E. and others. Effect of electric convulsive shock on human capillaries. Confinia neurol., **15**: 84–90, 1955.
6. Dazzi, I. and Basevi, D. Thrombocytopenic purpura following electroshock. Progr. med., Napoli, **13**: 105–109, 1957.
7. Fantl, P. Simon, S. E. Fibrinolysis following electrically induced convulsions. Australian J. Exper. Biol. & M. Sc., **26**: 521–529, 1948.
8. Lowinger, L. and Huddleson, J. H. Complications in electric shock therapy. Am. J. Psychiat., **102**: 594–598, 1946.
9. Williams, R. L. and Barrera, S. E. Observations and opinions concerning complications and contraindications in electric convulsive therapy. Psychiat. Quart., **24**: 800–809, 1950.